Genet. Sel. Evol. 37 (2005) 31–56

© INRA, EDP Sciences, 2004
DOI: 10.1051/gse:2004035

Original article

A general approach to mixed effects modeling of residual variances in generalized linear mixed models

Kadir KIZILKAYA*, Robert J. TEMPELMAN**

Department of Animal Science, Michigan State University, East Lansing, MI 48824-1225, USA

(Received 25 March 2004; accepted 24 July 2004)

Abstract – We propose a general Bayesian approach to heteroskedastic error modeling for generalized linear mixed models (GLMM) in which linked functions of conditional means and residual variances are specified as separate linear combinations of fixed and random effects. We focus on the linear mixed model (LMM) analysis of birth weight (BW) and the cumulative probit mixed model (CPMM) analysis of calving ease (CE). The deviance information criterion (DIC) was demonstrated to be useful in correctly choosing between homoskedastic and heteroskedastic error GLMM for both traits when data was generated according to a mixed model specification for both location parameters and residual variances. Heteroskedastic error LMM and CPMM were fitted, respectively, to BW and CE data on 8847 Italian Piemontese first parity dams in which residual variances were modeled as functions of fixed calf sex and random herd effects. The posterior mean residual variance for male calves was over 40% greater than that for female calves for both traits. Also, the posterior means of the standard deviation of the herd-specific variance ratios (relative to a unitary baseline) were estimated to be 0.60 ± 0.09 for BW and 0.74 ± 0.14 for CE. For both traits, the heteroskedastic error LMM and CPMM were chosen over their homoskedastic error counterparts based on DIC values.

Bayesian analysis / genetic evaluation / heterogeneous variances / threshold model

1. INTRODUCTION

Most genetic evaluation programs for livestock are based on generalized linear mixed model (GLMM) analyses [39] with breeding values or genetic effects modeled as random with a covariance structure based on known genetic relationships in a pedigree. Often heterogeneous residual variances, also

* Current address: Department of Animal Science, Adnan Menderes University, Aydin, 09100 Turkey.

** Corresponding author: tempelma@msu.edu

defined as residual heteroskedasticity, are observed across various environmental subclasses. If heteroskedasticity is not properly modeled, biased breeding value predictions may result such that a disproportionate numbers of animals are selected from highly variable environmental subclasses with consequent adverse effects on genetic improvement [16].

Genetic evaluation systems based on linear mixed models (LMM) have been proposed for the analysis of continuous production characters in which residual variances are modeled as a function of both fixed, *e.g.* region and sex, and random, *e.g.* herd, effects [9, 17, 27, 29, 31]. It is well recognized that using random effects or, equivalently, empirical Bayes specifications for a factor with many small subclasses facilitates the efficient borrowing of information from all subclasses for inference on any one particular subclass [13, 30]. There has been increasing interest to model residual heteroskedasticity in GLMM other than the LMM. For example, the cumulative probit mixed model (CPMM), often labeled as the threshold model in animal breeding [12], was developed to provide genetic evaluations on ordinal categorical traits, such as calving ease (CE) as the LMM is only suitable for the analysis of nearly normally distributed characters. The CPMM was extended by Foulley and Gianola [10] to model the logarithm of residual variances as a linear function of fixed effects on the underlying latent scale that characterizes this model. This heteroskedastic CPMM now forms the basis for CE genetic evaluations in French Montbéliarde, Normande, Holstein [7] and Italian Holstein [3] cattle by which residual variances are modeled, for example, as a function of calf sex and age of dam. San Cristobal-Gaudy *et al.* [32] further extended the CPMM of Foulley and Gianola [10] to specify log residual variances as a function of both fixed and random effects, including correlated genetic effects.

With the exception of Sorensen and Waagepetersen [37], virtually all of the heteroskedastic error GLMM analyses procedures presented thus far have involved either rather tenuous large sample approximations or complicated numerical integrations as necessary to derive marginal likelihood functions. Furthermore, we perceive the lack of a unifying framework for the structural modeling of heterogeneous variances in GLMM analysis, whether for LMM, CPMM, or other models such as those for censored data [36] or count data [40]. The objectives of our study then were (1) to develop and validate a fully Bayesian structural mixed effects multiplicative model for residual variances in a GLMM, concentrating on a LMM analysis of normally distributed data and a CPMM analysis of ordinal data and (2) to apply the model to a dataset of BW and CE observed on calvings from first parity Italian Piemontese dams.

2. MODEL CONSTRUCTION

For a number of GLMM, data augmentation schemes exist such that a $n \times 1$ vector of either observed or augmented variables $\mathbf{L} = \{L_i\}_{i=1}^n$ conceptually maps one-to-one to the data vector $\mathbf{Y} = \{Y_i\}_{i=1}^n$. Examples where such augmented variables have been particularly useful for Bayesian inference include the CPMM [35] and censored data models [36].

We write a linear mixed effects model as

$$\mathbf{L} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{1}$$

where $\boldsymbol{\beta}$ is a vector of fixed location effects, $\mathbf{u}$ is a vector of random location effects, $\mathbf{X}$ and $\mathbf{Z}$ are known design matrices and $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}(\boldsymbol{\xi}))$ is a vector of normally distributed residuals with variance covariance matrix $\mathbf{R}(\boldsymbol{\xi})$ having a certain heteroskedasticity specification based on unknown $\boldsymbol{\xi}$ as defined later. The linear model in (1) is equivalent to the following distributional specification:

$$\mathbf{L} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi} \sim p(\mathbf{L} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi}) = \mathcal{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{R}(\boldsymbol{\xi})). \tag{2}$$

For a LMM of normally distributed data, there is no distinction between $\mathbf{Y}$ and $\mathbf{L}$, *i.e.* $\mathbf{Y} \equiv \mathbf{L}$, such that $p(\mathbf{Y} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi}) \equiv p(\mathbf{L} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi})$. However, for ordinal data with say, $C = 4$ categories, $\mathbf{L}$ maps to $\mathbf{Y}$ as follows:

$$Y_i = \begin{cases} 1 & \text{if } \tau_0 < L_i \leq \tau_1, \\ 2 & \text{if } \tau_1 < L_i \leq \tau_2, \\ 3 & \text{if } \tau_2 < L_i \leq \tau_3, \text{ and} \\ 4 & \text{if } \tau_3 < L_i \leq \tau_4; \quad i = 1 \ldots n, \end{cases} \tag{3}$$

where $\tau_0 = -\infty < \tau_1 < \tau_2 < \tau_3 < \tau_4 = \infty$ are threshold parameters that define bin boundaries for $\mathbf{Y}$ based on $\mathbf{L}$.

We further partition $\mathbf{e}$ into residual variance subclasses $\mathbf{e}' = [\mathbf{e}'_{11}\ \mathbf{e}'_{12}\ \cdots\ \mathbf{e}'_{st}]$ where $\mathbf{e}_{kl} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{kl}(\boldsymbol{\xi}) = \mathbf{I}_{n_{kl}}\sigma^2_{e_{kl}})$ pertains to the $n_{kl} \times 1$ subvector of residuals identified with the kth level ($k = 1, 2, \ldots, s$) of a single fixed factor, *e.g.* sex, and the lth level ($l = 1, 2, \ldots, t$) of a single random factor, *e.g.* herd, that jointly influence the residual variance $\sigma^2_{e_{kl}}$ for the klth group. For simplicity of presentation, we concentrate on a specification based on only one fixed factor and one random factor for residual heteroskedasticity; implementation details involving multiple fixed and random factors are forthcoming in future work. We partition the data realization $\mathbf{y}$ of $\mathbf{Y}$ as $\mathbf{y} = [\mathbf{y}'_{11}\ \mathbf{y}'_{12}\ \cdots\ \mathbf{y}'_{st}]'$ and the corresponding augmented variables $\mathbf{L} = [\mathbf{L}'_{11}\ \mathbf{L}'_{12}\ \cdots\ \mathbf{L}'_{st}]'$ as with $\mathbf{e}$. We propose a

multiplicative structural effects model as follows:

$$\sigma^2_{e_{kl}} = \sigma^2_e \gamma_k \nu_l; \quad k = 1, 2, \ldots, s; \quad l = 1, 2, \ldots, t; \tag{4}$$

where σ^2_e is a "fixed" reference residual variance, $\gamma_k > 0$ is a multiplicative scaling factor identified with the kth level of the fixed effect subclass and $\nu_l > 0$ is a random multiplicative scaling factor unique to the lth level of the random effect. When noninformative priors are used for σ^2_e and $\boldsymbol{\gamma}$, identifiability constraints are required. In that case, we arbitrarily set $\gamma_s = 1$ such that σ^2_e then specifies the expected residual variance for the last fixed effects subclass $k = s$ given $\mathrm{E}(\nu_l) = 1$. If one writes equation (4) on the logarithmic scale, this constraint ($\log(\gamma_s) = 0$) is analogous to the corner parameterization used for location parameters [5] and is the default parameterization using, for example, SAS® linear model software [21].

Note then that fixed and random effects specifications are specified for both location parameters, *i.e.* $\boldsymbol{\beta}$ and $\mathbf{u}$, and for dispersion parameters, *i.e.* σ^2_e, $\boldsymbol{\gamma} = [\gamma_1\ \gamma_2\ \ldots\ \gamma_s]'$ and $\mathbf{v} = [\nu_1\ \nu_2\ \ldots\ \nu_t]'$. Different classes of effects could be specified for the two sets of parameters. For example, for fixed effects, there may be calf sex differences specified for means, *i.e.* location parameters, but not for residual variances whereas for random effects there may be variability in contemporary groups specified for residual variances but not for location parameters.

We specify a subjective prior density on $\boldsymbol{\beta}$:

$$\boldsymbol{\beta} \sim p(\boldsymbol{\beta}) \tag{5}$$

where $p(\boldsymbol{\beta})$ is typically specified to be a bounded flat uniform prior or an informative, perhaps conjugate normal, prior. Here $\boldsymbol{\beta}$ could be parameterized in a number of different ways, *e.g.* to partition factor level effects and interactions thereof from a reference mean, say μ, analogous to the partitioning of σ^2_e from $\boldsymbol{\gamma}$ for residual variance modeling. For elements of $\boldsymbol{\beta}$, as with $\boldsymbol{\gamma}$, identifiability constraints should be specified if non-informative priors are used [5, 21].

As typical for any GLMM, $\mathbf{u}$ is specified by a structural multivariate prior:

$$\mathbf{u} \,|\, \boldsymbol{\varphi} \sim p(\mathbf{u} \,|\, \boldsymbol{\varphi}) = \mathcal{N}(\mathbf{0}, \mathbf{G}(\boldsymbol{\varphi})). \tag{6}$$

Here $\mathbf{G}(\boldsymbol{\varphi})$ is a variance-covariance matrix that is a function of several unknown variance components or variance-covariance matrices in $\boldsymbol{\varphi}$, depending on whether or not there are multiple sets of random effects and/or specified covariances between these sets; an example of the latter is the covariance between additive and maternal genetic effects. Furthermore, flat priors, inverted

Gamma densities, inverted Wishart densities or products of conditionally independent components thereof may be specified for the prior density on $\boldsymbol{\varphi}$, *i.e.*

$$\boldsymbol{\varphi} \sim p(\boldsymbol{\varphi}) \tag{7}$$

depending, again, on the number of sets of random effects and whether there are any covariances thereof [34].

Analogously to $\boldsymbol{\beta}$, a flat (bounded) prior density may be specified separately for σ_e^2 and for each γ_k; alternatively, proper conjugate inverted-gamma or other subjective priors may be specified, *i.e.*

$$\sigma_e^2 \sim p(\sigma_e^2) \tag{8a}$$

and

$$\gamma_k \sim p(\gamma_k), \quad k = 1, 2, \ldots, s. \tag{8b}$$

Conversely, a structural prior is used to model the random residual dispersion effects, ν_l, $l = 1, 2, \ldots, t$, analogous to that used for the random location effects $\mathbf{u}$ in equation (6). We conveniently choose this structural prior to be an inverted-gamma density with parameters α_ν and $\alpha_\nu - 1$:

$$p(\nu_l \mid \alpha_\nu) = \frac{(\alpha_\nu - 1)^{\alpha_\nu}}{\Gamma(\alpha_\nu)} (\nu_l)^{-(\alpha_\nu+1)} \exp\left(-\frac{(\alpha_\nu - 1)}{\nu_l}\right); \; l = 1, 2, \ldots, t; \; \alpha_\nu > 2. \tag{9}$$

Here $\mathrm{E}(\nu_l) = 1$ and $\sigma_\nu^2 = \mathrm{Var}(\nu_l) = \frac{1}{\alpha_\nu - 2}$, $l = 1, 2, \ldots, t$. Note that as $\alpha_\nu \to \infty$ then $\sigma_\nu^2 \to 0$ such that the influence of $\mathbf{v}$ on residual heteroskedasticity diminishes. The structural prior in (9) facilitates a borrowing of information across the t random effects of $\mathbf{v}$, similar to what occurs for classical random effects modeling of location parameters using $\mathbf{u}$ [30]. Since $\mathrm{E}(\nu_l) = 1$, ν_l can be interpreted as a relative measure of variance for random subclass l. Generally, α_ν is unknown such that a subjective prior may be placed on it. One vaguely informative proper prior commonly used for strictly positive parameters [1] and used in some of our previous applications [20, 41] is specified by

$$\alpha_\nu \sim p(\alpha_\nu) \propto \frac{1}{(1 + \alpha_\nu)^2}. \tag{10}$$

The remaining hierarchical specifications in this heteroskedastic GLMM depend upon the first (data sampling) stage of the $n \times 1$ data vector $\mathbf{y}$. For a LMM analysis of normal error data, equation (1) suffices, *i.e.* $\mathbf{Y} \equiv \mathbf{L}$, such that no augmented variables are required. However, for a CPMM analysis of ordinal data

with C ordinal categories, numbered $j = 1, 2, \ldots, C$, we specify the first stage of our hierarchical model using Sorensen *et al*. [35]:

$$p(\mathbf{y} \mid \mathbf{L}, \boldsymbol{\tau}) = \prod_{k=1}^{s} \prod_{l=1}^{t} \prod_{i=1}^{n_{kl}} \left\{ \sum_{j=1}^{C} 1(\tau_{j-1} < L_{ikl} < \tau_j) 1(y_{ikl} = j) \right\}, \tag{11a}$$

where L_{ikl} and y_{ikl} are the ith elements of $\mathbf{L}_{kl}$ and $\mathbf{y}_{kl}$, respectively. As before, $\boldsymbol{\tau} = [\tau_0\, \tau_1\, \cdots\, \tau_C]'$ denotes a vector of unknown threshold parameters that delimit the augmented variables $\mathbf{L}$ into their respective observed data bins $\mathbf{y}$ with 1(.) denoting the indicator function having value 1 if the condition within the function is true and 0 otherwise.

The second stage of the model was given previously in equation (2) and can be re-expressed as a product of conditionally independent normals:

$$\mathbf{L} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\xi} \sim \prod_{k=1}^{s} \prod_{l=1}^{t} \prod_{i=1}^{n_{kl}} p(L_{ikl} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \gamma_k, \nu_l) \tag{11b}$$

with $\boldsymbol{\xi} = [\sigma_e^2\ \boldsymbol{\gamma}'\ \mathbf{v}']'$ and $p(L_{ikl} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \gamma_k, \nu_l) = \mathcal{N}(\mathbf{x}'_{ikl}\boldsymbol{\beta} + \mathbf{z}'_{ikl}\mathbf{u}, \sigma_e^2 \gamma_k \nu_l)$. Here, $\mathbf{x}'_{ikl}$ and $\mathbf{z}'_{ikl}$ are known incidence row vectors from $\mathbf{X}$ and $\mathbf{Z}$, respectively, corresponding to animal i, $i = 1, 2, \ldots, n_{kl}$, within the klth residual variance subclass, $k = 1, 2, \ldots, s; l = 1, 2, \ldots, t$.

Recall that $\tau_0 = -\infty$ and $\tau_C = +\infty$; furthermore, τ_1 is fixed to an arbitrary constant in order to satisfy identifiability constraints. We adapt the alternative parameterization presented by Sorensen *et al.* [35] in which the residual variance is explicitly modeled rather than being constrained, *e.g.* $\sigma_e^2 = 1$, the latter being the more common parameterization used in homoskedastic error CPMM [12]. This specification thereby requires one additional constraint on $\boldsymbol{\tau}$, say on $\tau_2 > \tau_1$, such that $C - 3$ parameters in $\boldsymbol{\tau}$ are uniquely identifiable. A prior distribution on the remaining elements of $\boldsymbol{\tau}$ may be specified provided that the order constraints on elements of $\boldsymbol{\tau}$ are satisfied [35], *i.e.*

$$\boldsymbol{\tau} = [\tau_3\ \tau_4\ \cdots\ \tau_{C-1}]' \sim p(\boldsymbol{\tau});\ \tau_3 < \cdots < \tau_{C-1}. \tag{12}$$

The joint posterior density of $\boldsymbol{\beta}$, $\mathbf{u}$, $\boldsymbol{\gamma}$, $\mathbf{v}$, $\boldsymbol{\varphi}$, α_ν and any other parameters necessary for the GLMM in question, *i.e.* $\mathbf{L}$ and $\boldsymbol{\tau}$ in the CPMM, is simply specified as the product of the various stages of the hierarchical model. That is, for the LMM where $\mathbf{Y} \equiv \mathbf{L}$, the joint posterior density of all unknowns

specified to proportionality is

$$\begin{aligned} &p(\boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}, \boldsymbol{\varphi}, \alpha_\nu \mid \mathbf{y}) \\ &\propto p(\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}) p(\boldsymbol{\beta}) p(\mathbf{u} \mid \boldsymbol{\varphi}) p(\boldsymbol{\varphi}) \\ &p(\sigma_e^2) \left(\prod_{k=1}^{s} p(\gamma_k) \right) \left(\prod_{l=1}^{t} p(\nu_l \mid \alpha_\nu) \right) p(\alpha_\nu) \end{aligned} \tag{13}$$

whereas for the CPMM, the joint posterior density is

$$\begin{aligned} &p(\mathbf{L}, \boldsymbol{\tau}, \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}, \boldsymbol{\varphi}, \alpha_\nu \mid \mathbf{y}) \\ &\propto p(\mathbf{y} \mid \mathbf{L}, \boldsymbol{\tau})\, p(\mathbf{L} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}) \\ &p(\boldsymbol{\tau}) p(\boldsymbol{\beta}) p(\mathbf{u} \mid \boldsymbol{\varphi}) p(\boldsymbol{\varphi}) p(\sigma_e^2) \left(\prod_{k=1}^{s} p(\gamma_k) \right) \left(\prod_{l=1}^{t} p(\nu_l \mid \alpha_\nu) \right) p(\alpha_\nu). \end{aligned} \tag{14}$$

A Markov chain Monte Carlo (MCMC) inference strategy requires determination of and sampling from the full conditional distributions (FCD) of each parameter or blocks thereof. The joint FCD for $\boldsymbol{\beta}$ and $\mathbf{u}$, regardless of whether inference is based on the LMM or the CPMM, can be readily shown to be multivariate normal

$$\begin{aligned} &\boldsymbol{\beta}, \mathbf{u} \mid \boldsymbol{\varphi}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}, \alpha_\nu, \mathbf{y}, \mathbf{L} \\ &\sim \mathcal{N}\left(\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix}, \begin{bmatrix} \mathbf{X}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{X} & \mathbf{X}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{Z} \\ \mathbf{Z}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{X} & \mathbf{Z}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{Z} + (\mathbf{G}(\boldsymbol{\varphi}))^{-1} \end{bmatrix}^{-1} \right) \end{aligned} \tag{15}$$

where $\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{X} & \mathbf{X}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{Z} \\ \mathbf{Z}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{X} & \mathbf{Z}'(\mathbf{R}(\boldsymbol{\xi}))^{-1}\mathbf{Z} + (\mathbf{G}(\boldsymbol{\varphi}))^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\ \mathbf{R}^{-1}\ \mathbf{L} \\ \mathbf{Z}'\ \mathbf{R}^{-1}\ \mathbf{L} \end{bmatrix}$ are the typical mixed model solutions to $\boldsymbol{\beta}$ and $\mathbf{u}$ based on current MCMC samples of other unknown parameters/variables. Furthermore, $(\mathbf{R}(\boldsymbol{\xi}))^{-1} = \bigoplus_{k=1,l=1}^{k=s,l=t} (\mathbf{R}(\boldsymbol{\xi})_{kl})^{-1}$ where $\bigoplus_{k=1,l=1}^{k=s,l=t}$ represents the direct sum operator [33] across all *st* groups. Univariate MCMC sampling strategies, as an alternative to a joint block sample from (15), are elucidated in Wang *et al.* [43].

As in Kizilkaya *et al*. [20], we use an efficient Metropolis Hastings update [6] to sample from the joint FCD of $\mathbf{L}$ and $\boldsymbol{\tau}$ under the CPMM. If some partitions of $\boldsymbol{\varphi}$ form a variance-covariance matrix, then their respective FCD can be readily shown to be inverted-Wishart whereas if other partitions of $\boldsymbol{\varphi}$ involve scalar variance but no covariance components, then the FCD of each component can be shown to be inverted-gamma [34].

We subsequently determine the FCD for σ_e^2 which is virtually identical for both the CPMM and the LMM. Under the CPMM,

$$p\left(\sigma_e^2 \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\gamma}, \mathbf{v}, \boldsymbol{\varphi}, \boldsymbol{\tau}, \alpha_\nu, \mathbf{L}, \mathbf{y}\right) \propto \left(\prod_{k=1}^{s}\prod_{l=1}^{t}\left(\sigma_e^2\gamma_k\nu_l\right)^{-\frac{n_{kl}}{2}}\right)$$
$$\times \exp\left(-\frac{1}{2}\sum_{k=1}^{s}\sum_{l=1}^{t}\frac{\left(\mathbf{L}_{kl}-\mathbf{X}_{kl}\boldsymbol{\beta}-\mathbf{Z}_{kl}\mathbf{u}\right)'\left(\mathbf{L}_{kl}-\mathbf{X}_{kl}\boldsymbol{\beta}-\mathbf{Z}_{kl}\mathbf{u}\right)}{\sigma_e^2\gamma_k\nu_l}\right)p\left(\sigma_e^2\right)$$
$$\propto \left(\left(\sigma_e^2\right)^{-\frac{n_{++}}{2}}\right)\exp\left(-\frac{\sum_{k=1}^{s}\sum_{l=1}^{t}\frac{\mathbf{e}_{kl}'\mathbf{e}_{kl}}{2\gamma_k\nu_l}}{\sigma_e^2}\right)p\left(\sigma_e^2\right) \quad (16)$$

where $\mathbf{X}_{kl}$ and $\mathbf{Z}_{kl}$ are appropriate partitions of $\mathbf{X}$ and $\mathbf{Z}$ that relate $\mathbf{y}_{kl}$ and $\mathbf{L}_{kl}$ to $\boldsymbol{\beta}$ and $\mathbf{u}$, respectively. With a flat, preferably bounded, uniform prior specified in (8a), *i.e.* $p(\sigma_e^2) \propto 1$, the FCD for σ_e^2 from (16) can then be seen to be simply inverted gamma with parameters $\alpha_{++} = \frac{n_{++}}{2} - 1$ and $\beta_{++} = \frac{1}{2}\sum_{k=1}^{s}\sum_{l=1}^{t}\frac{\mathbf{e}_{kl}{}'\mathbf{e}_{kl}}{\gamma_k\nu_l}$, where $n_{++} = \sum_{k=1}^{s}\sum_{l=1}^{t} n_{kl}$. Alternatively, if one specifies (8a) to be any proper inverted gamma density, then the FCD of σ_e^2 will also be inverted-gamma.

It can also be readily determined that the FCD of γ_k is inverted gamma with parameters $\alpha_{k+} = \frac{\sum_{l=1}^{t} n_{kl}}{2} - 1$ and $\beta_{k+} = \frac{1}{2\sigma_e^2}\sum_{l=1}^{t}\frac{\mathbf{e}_{kl}'\mathbf{e}_{kl}}{\nu_l}$, $k = 1, 2, \ldots, s-1$ based on a flat uniform prior specification for (8b), whereas the FCD of ν_l is inverted gamma with parameters $\alpha_{+l} = \frac{\sum_{k=1}^{s} n_{kl}}{2} + \alpha_\nu$ and $\beta_{+l} = \frac{1}{2\sigma_e^2}\sum_{k=1}^{s}\frac{\mathbf{e}_{kl}'\mathbf{e}_{kl}}{\gamma_k} + \alpha_\nu - 1$ for $l = 1, 2, \ldots, t$. Note that MCMC sampling of elements of $\boldsymbol{\xi}$ under the LMM is identical to that for the CPMM except that $\mathbf{e}_{kl} = \mathbf{y}_{kl} - \mathbf{X}_{kl}\boldsymbol{\beta} - \mathbf{Z}_{kl}\mathbf{u}$ rather than $\mathbf{e}_{kl} = \mathbf{L}_{kl} - \mathbf{X}_{kl}\boldsymbol{\beta} - \mathbf{Z}_{kl}\mathbf{u}$.

The FCD for α_ν based on the prior $p(\alpha_\nu)$ adopted from equation (10), is not readily recognizable:

$$p(\alpha_\nu \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\varphi}, \boldsymbol{\gamma}, \mathbf{v}, \mathbf{y}, \mathbf{L}, \boldsymbol{\tau}) \propto \frac{(\alpha_\nu - 1)^{\alpha_\nu t}}{(\Gamma(\alpha_\nu))^t} \exp\left(-(\alpha_\nu - 1)\sum_{l=1}^{t} \nu_l^{-1}\right) \prod_{l=1}^{t} (\nu_l)^{-(\alpha_\nu+1)}\, p(\alpha_\nu). \tag{17}$$

Here we sample from (17) indirectly using the change of variable $\psi_\nu = \log(\alpha_\nu)$ with a random walk Metropolis-Hastings update based on a Gaussian proposal density [4] in a manner virtually identical to how we sampled the degrees of freedom parameter for a cumulative t-link mixed model (CTMM) in our previous work [20].

3. MODEL COMPARISON

The deviance information criterion (DIC) has been proposed for comparing goodness of fit for alternative constructions of hierarchical models to data [38]. The DIC is based on the posterior distribution of the deviance statistic or -2 times the sampling distribution of the data as specified in the first stage of a hierarchical model. For the LMM, the data sampling stage is specified as:

$$p\left(\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}\right) = \prod_{k=1}^{s}\prod_{l=1}^{t} (2\pi)^{-\frac{n_{kl}}{2}} (\sigma_e^2 \gamma_k \nu_l)^{-\frac{n_{kl}}{2}} \exp\left(-\frac{[\mathbf{y}_{kl} - \mathbf{X}_{kl}\boldsymbol{\beta} - \mathbf{Z}_{kl}\mathbf{u}]'[\mathbf{y}_{kl} - \mathbf{X}_{kl}\boldsymbol{\beta} - \mathbf{Z}_{kl}\mathbf{u}]}{2\sigma_e^2 \gamma_k \nu_l}\right) \tag{18}$$

whereas for the CPMM, the data sampling stage is based on the marginal likelihood of the data [20] based on the integration of the product of (11a) and (11b) jointly across all n observations with respect to $\mathbf{L}$:

$$\begin{aligned} p\left(\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{u}, \sigma_e^2, \boldsymbol{\gamma}, \mathbf{v}\right) &= \prod_{k=1}^{s}\prod_{l=1}^{t}\prod_{i=1}^{n_{kl}} \mathrm{Prob}\left(Y_{ikl} = y_{ikl} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\tau}, \sigma_e^2, \gamma_k, \nu_l\right) \\ &= \prod_{k=1}^{s}\prod_{l=1}^{t}\prod_{i=1}^{n_{kl}} \left(\Phi\left(\frac{\tau_{y_{ikl}} - \left(\mathbf{x}_{ikl}'\boldsymbol{\beta} + \mathbf{z}_{ikl}'\mathbf{u}\right)}{\sqrt{\sigma_e^2 \gamma_k \nu_l}}\right) - \Phi\left(\frac{\tau_{y_{ikl}} - \left(\mathbf{x}_{ikl}'\boldsymbol{\beta} + \mathbf{z}_{ikl}'\mathbf{u}\right)}{\sqrt{\sigma_e^2 \gamma_k \nu_l}}\right)\right). \end{aligned} \tag{19}$$

Smaller DIC values are indicative of better data fit. Further details on DIC including its components and computations for LMM and CPMM can be found in Kizilkaya *et al.* [20].

4. DATA

4.1. Simulation study

A simulation study was carried out to validate Bayesian inference on the proposed heteroskedastic error LMM and CPMM and to assess the ability of the DIC to correctly choose between homoskedastic and heteroskedastic error GLMM. A simple mixed effects model was used to generate latent variables $\mathbf{L}$ for 50 progeny from each of 50 unrelated sires:

$$L_{ijkl} = \mu + sex_i + herd_j + sire_k + e_{ijkl}.$$

Here $\mu = 0.5$, $\{sex_i\}_{i=1}^{2}$ ($sex_1 = -0.5$ and $sex_2 = 0.5$) represents a 2×1 vector of fixed sex effects. Furthermore, $\{herd_j\}_{j=1}^{100} \sim N(\mathbf{0}, \mathbf{I}\,\sigma_h^2)$ is a 100×1 vector of independent random location effects, and $\{sire_k\}_{k=1}^{50} \sim N(\mathbf{0}, \mathbf{I}\sigma_s^2)$ represent a 50×1 vector of independent random location effects with $\sigma_h^2 = 0.25$ and $\sigma_s^2 = 0.10$. Hence, L_{ijkl} and e_{ijkl} are, respectively, the latent and residual variables generated for calf l of sex i with sire k and raised within herd j. Now, $\mathbf{e}_{ij} = \{e_{ijkl}\} \sim N(\mathbf{0}, \mathbf{I}_{n_{ij}}\sigma_e^2\gamma_i\nu_j)$ is the vector of residuals associated with the n_{ij} records from sex i and herd j where $\nu_j \sim$ Inverted-Gamma$(\alpha_\nu, \alpha_\nu - 1)$, $j = 1, 2, \ldots, 100$. Three replicated datasets from each of four different populations or different values of α_ν were generated. Three of the populations had residual heteroskedasticity specifications that differed with respect to degree of herd variability: Population I: $\alpha_\nu = 3$, Population II: $\alpha_\nu = 12$, and Population III: $\alpha_\nu = 50$. That is, the values $\alpha_\nu = 3$, 12 and 50 represented extreme, moderate and mild levels, respectively, of residual heteroskedasticity across herds, given that the standard deviations $\sigma_\nu = \sqrt{\frac{1}{\alpha_\nu - 2}}$ of the relative variance factors are then 1, $\sqrt{\frac{1}{10}}$, and $\sqrt{\frac{1}{48}}$, respectively. Each population was based on $\sigma_e^2 = 1.25$ and $\gamma_1 = 0.80$, with constraint $\gamma_2 = 1$ meaning that Sex 1 (say females) had a marginal residual variance of $\sigma_{e_{female}}^2 = \sigma_e^2\gamma_1 = 1.00$ whereas males (Sex 2) had a marginal residual variance of $\sigma_{e_{male}}^2 = \sigma_e^2\gamma_2 = 1.25$. Our MCMC implementation was slightly reparameterized from that presented previously in that $\sigma_{e_{male}}^2$ and $\sigma_{e_{female}}^2$ were sampled directly. The three replicates from Population IV were generated with homoskedastic error ($\alpha_\nu = \infty$) with $\sigma_e^2 = 1$ and $\gamma_1 = \gamma_2 = 1$.

Levels of fixed and random effects for both location and residual variance parameters were randomly assigned to individuals in generating data. Augmented data $\mathbf{L}$ was mapped to ordinal data $\mathbf{y}$ based on $C = 4$ categories with thresholds $\tau_1 = -0.50$, $\tau_2 = 1.00$ and $\tau_3 = 2.00$ in all populations. Here $\mathbf{L}$ and

its ordinal mapping to **y** were analyzed using the appropriate GLMM (LMM and CPMM, respectively) based on both homoskedastic and heteroskedastic error models with flat priors placed on $\boldsymbol{\beta}$, $\sigma^2_{e_{male}}$, $\sigma^2_{e_{female}}$, σ^2_s, σ^2_h, and the prior from equation (10), used for α_ν. We invoked the constraints $\tau_1 = -0.50$ and $\tau_2 = 1.00$ to facilitate parameter identifiability in the CPMM analysis of ordinal **y**. Although these constraints were identical to the true values of τ_1 and τ_2, so as to avoid rescaling of estimates of other parameters for comparison to their true values, any values of τ_1 and τ_2 could have been used provided $\tau_2 > \tau_1$. For each replicated data set within each population, a burn-in period of 10 000 MCMC cycles was discarded before saving samples from each of an additional 100 000 MCMC cycles. DIC values were computed for each model on each replicated dataset to validate those measures as appropriate for model choice. Posterior densities on $\sigma^2_{e_{female}}$, $\sigma^2_{e_{male}}$, and α_ν were summarized by their 95% equal-tailed posterior probability intervals (95% PPI), *i.e.* the 2.5th and 97.5th percentiles of the posterior densities.

4.2. Italian Piemontese birth weight and calving ease data

The very same first parity calf BW and CE scores analyzed by Kizilkaya *et al.* [20] are also considered here. These data were recorded on Italian Piemontese cattle from January, 1989 to July, 1998 by Associazione Nazionale Allevatori Bovini di Razza Piemontese (ANABORAPI), Strada Trinità 32a, 12061 Carrù, Italy. Only herds that were represented by at least 100 records were considered in the study, leaving a total of 8847 records. CE was scored into five categories by breeders and subsequently recorded by technicians who visited the breeders monthly. The five ordered categories were: (1) unassisted delivery, (2) assisted easy calving, (3) assisted difficult calving, (4) Caesarean section, and (5) foetotomy. As the incidence of foetotomy was less than 0.5%, the last two ordinal categories were combined, leaving a total of four mutually exclusive categories. The effects of dam age, sex of the calf, and their interaction were modeled by combining eight different age groups with sex of calf for a total of 16 nominal subclasses. Herd-year-season (HYS) subclasses were created from combinations of herd, year, and two different seasons (from November to April and from May to October). Further details on the data, including factor definitions and number of levels thereof, are provided in Kizilkaya *et al.* [20].

As in Kizilkaya *et al.* [19,20], the LMM and CPMM used for the analyses of BW and CE included the fixed effects of age of dam classifications, sex of calf and their interaction in $\boldsymbol{\beta}$, whereas the random effects **u** included independent

herd-year-season effects in **h**, random sire effects in **s** and random maternal grandsire effects in **m**. We also assume:

$$\begin{pmatrix}\mathbf{s}\\ \mathbf{m}\end{pmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \mathbf{G} = \mathbf{G}_0 \otimes \mathbf{A}\right)$$

and

$$\mathbf{h} \sim N\left(\mathbf{0}, \mathbf{I}\sigma_h^2\right)$$

where $\mathbf{G}_0 = \begin{pmatrix}\sigma_s^2 & \sigma_{sm}\\ \sigma_{sm} & \sigma_m^2\end{pmatrix}$, with σ_s^2 denoting the sire variance, σ_m^2 denoting the maternal grandsire variance, σ_{sm} denoting the sire-maternal grandsire covariance, and σ_h^2 denoting the HYS variance. Furthermore, $\otimes$ denotes the Kronecker product [33], and **A** is the numerator additive relationship matrix between sires due to identified male ancestors [14].

Residual heteroskedasticity was modeled as a function of fixed sex effects and random herd effects. For fixed effects, $\gamma_{male} = 1$ was considered as a constraint such that $\sigma^2_{e_{female}} = \sigma_e^2\gamma_{female}$ is the marginal residual variance for female calves and $\sigma^2_{e_{male}} = \sigma_e^2$ is the marginal residual variance for male calves as specified in the simulation study. Random effects of 66 herds for residual variability were specified by $\nu_j \sim$ Inverted-Gamma$(\alpha_\nu, \alpha_\nu - 1)$, $j = 1, 2, \ldots, 66$. Flat (bounded) uniform priors were placed on $\boldsymbol{\beta}$, σ_s^2, σ_m^2, σ_{sm}, σ_h^2, $\sigma^2_{e_{male}}$, $\sigma^2_{e_{female}}$, and the noninformative prior from equation (10) was used for α_ν.

MCMC inference was based on the running of three independent chains for each model using the same specifications as in Kizilkaya *et al.* [20]. That is, for each chain, a total of 20 000 burn-in MCMC cycles was followed by the saving of samples from each of 100 000 additional MCMC cycles.

Key genetic parameters were calf sex specific, including direct heritabilities, *i.e.* $h^2_{d_female}$ and $h^2_{d_male}$, maternal heritabilities, *i.e.* $h^2_{m_female}$ and $h^2_{m_male}$, as based on the same expressions used by Kizilkaya *et al.* [19, 20] except that $\sigma^2_{e_{male}}$ or $\sigma^2_{e_{female}}$ was substituted for σ_e^2 in the denominator for males and female calves, respectively. Posterior densities of these parameters, $\sigma_\nu = \frac{1}{\sqrt{\alpha_\nu - 2}}$, all variance components, the direct-maternal genetic correlation (r_{dm}) and key differences such as $h^2_{d_female} - h^2_{d_male}$, $h^2_{m_female} - h^2_{m_male}$, and $\sigma^2_{e_{male}} - \sigma^2_{e_{female}}$ were summarized by posterior means and standard deviations well as by 95% PPI. For computational expediency, 95% PPI on the herd specific relative variance measures ν_j, $j = 1, 2, \ldots, 66$, were approximated by their posterior means $\pm$ two posterior standard deviations. Analyses of both BW and CE were also carried out by the appropriate homoskedastic GLMM in order to facilitate DIC

comparisons against their heteroskedastic counterparts for model fit. Furthermore, Spearman rank correlations were computed between the heteroskedastic and homoskedastic GLMM for posterior means of **s** for both LMM and CPMM analyses of BW and CE, respectively.

5. RESULTS

5.1. Simulation study

Posterior inference on σ_s^2 and σ_h^2 is not reported in detail but summarized as follows. Posterior means of σ_s^2 and σ_h^2 were generally slightly biased upwards for the LMM analysis on **L** and for the CPMM analysis on **y**. Nevertheless, the 95% PPI of these two variance components, whether for the LMM analysis of **L** or the CPMM analysis of **y**, included the true values of parameters in all cases. The 95% PPI for these two components were also understandably wider for the CPMM analysis of **y** than for the LMM analysis of **L**, likely due to the substantial loss in data information in mapping from continuous **L** to ordinally discrete **y**.

The 95% PPI on $\sigma_{e_{male}}^2$, $\sigma_{e_{female}}^2$ and α_ν for each of the three replicated datasets from each of the four populations, are provided in Table I for the direct LMM analysis of **L** and Table II for CPMM for the corresponding ordinally mapped values of **y**. The 95% PPI of $\sigma_{e_{female}}^2$ and $\sigma_{e_{male}}^2$ in either the LMM and CPMM analyses of **L** and **y**, respectively, always included the true values of $\sigma_{e_{female}}^2 = 1$ and $\sigma_{e_{male}}^2 = 1.25$, respectively, for each of the three replicated datasets from each of Populations I, II, and III. For Population IV, in which three replicates were generated with homoskedastic error, posterior means and 95% PPI for $\sigma_{e_{female}}^2$ and $\sigma_{e_{male}}^2$ were found to be similar to each other and concentrated around the true value $\sigma_{e_{female}}^2 = \sigma_{e_{male}}^2 = 1$ as anticipated. For all populations, the 95% PPI of $\sigma_{e_{female}}^2$ and $\sigma_{e_{male}}^2$ tended to center about their corresponding posterior means (results not shown). The 95% PPI for α_ν for each of the three datasets from each of Populations I, II, and III also included the corresponding true parameter value. The widths of the 95% PPI of α_ν increased appreciably as α_ν increased. Again, because of the lower information content of ordinal data relative to underlying latent data, CPMM analyses of **y** produced understandably wider 95% PPI on α_ν relative to LMM analyses of **L**. Furthermore, posterior means of α_ν were biased upwards for both analyses in Population III (results not shown), thereby reflecting the greater skewness of the posterior densities for higher values of α_ν. As anticipated, the 95% PPI of α_ν in Population IV ($\alpha_\nu = \infty$) included very large values (>300) for either the LMM or CPMM analyses.

Table I. The 95% equal-tailed posterior probability intervals [2.5th percentile; 97.5th percentile] on residual dispersion parameters for four simulated populations with different levels of residual heteroskedasticity.

Population-Datasets[b]	LMM analysis[a]			CPMM analysis[a]		
	$\sigma^2_{e_{female}}$	$\sigma^2_{e_{male}}$	α_ν	$\sigma^2_{e_{female}}$	$\sigma^2_{e_{male}}$	α_ν
I-1	[0.76; 1.05]	[0.93; 1.29]	[2.53; 4.33]	[0.78; 1.14]	[0.95; 1.41]	[2.59; 5.08]
I-2	[0.87; 1.20]	[1.07; 1.47]	[2.53; 4.24]	[0.85; 1.24]	[0.92; 1.36]	[2.55; 4.84]
I-3	[0.78; 1.08]	[1.01; 1.39]	[2.59; 4.37]	[0.74; 1.11]	[0.70; 1.27]	[3.02; 6.53]
II-1	[0.95; 1.18]	[1.13; 1.41]	[5.65; 12.93]	[0.71; 0.99]	[0.88; 1.32]	[4.61; 13.47]
II-2	[0.84; 1.05]	[1.11; 1.38]	[5.94; 14.11]	[0.83; 1.24]	[1.13; 2.11]	[5.65; 21.28]
II-3	[0.92; 1.13]	[1.22; 1.49]	[8.43; 26.67]	[0.88; 1.24]	[1.05; 1.65]	[8.90; 182.6]
III-1	[0.88; 1.04]	[1.16; 1.38]	[21.2; 723.7]	[0.87; 1.12]	[1.15; 1.53]	[13.9; 2459]
III-2	[0.99; 1.18]	[1.32; 1.11]	[19.0; 439.3]	[0.91; 1.20]	[1.07; 1.44]	[10.2; 534.8]
III-3	[0.93; 1.12]	[1.16; 1.39]	[17.2; 248.2]	[0.92; 1.27]	[1.25; 1.89]	[9.78; 481.5]
IV-1	[0.92; 1.09]	[0.90; 1.06]	[33.2; 3317]	[0.89; 1.15]	[0.87; 1.14]	[24.9; 3849]
IV-2	[0.98; 1.16]	[0.97; 1.14]	[38.6; 3809]	[0.93; 1.22]	[0.96; 1.29]	[9.68; 457.7]
IV-3	[0.92; 1.09]	[0.91; 1.09]	[23.0; 1268]	[0.95; 1.42]	[1.03; 1.91]	[10.1; 362.3]

[a] Analyses included a linear mixed effects model (LMM) analysis of normally distributed data and cumulative probit mixed effects model (CPMM) analysis of ordinal data based on four response categories as mapped from the normally distributed data.

[b] First value indicates population (I–IV) and second value indicates replicate dataset (1–3) within each population. Residual variances were specified to be heterogeneous within the first three populations ($\sigma^2_{e_{female}} = 1.00$; $\sigma^2_{e_{male}} = 1.25$) with large ($\alpha_\nu = 3$), moderate ($\alpha_\nu = 12$) and small ($\alpha_\nu = 50$) levels of residual heteroskedasticity across random herd effects in Populations I, II, and III, respectively. Residual homoskedasticity was specified for all three datasets in Population IV ($\sigma^2_{e_{female}} = \sigma^2_{e_{male}} = 1$; $\alpha_\nu = \infty$).

Table II. Posterior inference on variance components allowing for residual heteroskedasticity based on a linear mixed model analysis of birth weight and a cumulative probit mixed model analysis calving ease scores in Italian Piemontese cattle.

	Linear mixed model analysis			Cumulative probit mixed model analysis		
Parameter[a]	PM ± SD[b]	95% PPI[c]	ESS[d]	PM ± SD[b]	95% PPI[c]	ESS[d]
σ^2_s	1.13 ± 0.20	[0.77; 1.54]	3379	0.13 ± 0.02	[0.09; 0.18]	2894
σ^2_m	0.50 ± 0.11	[0.31; 0.74]	1836	0.02 ± 0.01	[0.01; 0.04]	841
σ_{sm}	0.35 ± 0.11	[0.15; 0.56]	2803	0.02 ± 0.01	[0.003; 0.05]	1169
σ^2_h	1.68 ± 0.19	[1.33; 2.07]	14 094	0.13 ± 0.02	[0.10; 0.16]	7921
$\sigma^2_{e_{male}}$	14.44 ± 1.03	[12.63; 16.70]	7794	1.09 ± 0.09	[0.93; 1.29]	5915
$\sigma^2_{e_{female}}$	10.19 ± 0.73	[8.89; 11.77]	7713	0.71 ± 0.06	[0.61; 0.84]	5856
$\sigma^2_{e_{female}} - \sigma^2_{e_{male}}$	4.26 ± 0.53	[3.29; 5.36]	21 839	0.38 ± 0.05	[0.28; 0.49]	14 045
σ_ν	0.60 ± 0.09	[0.46; 082]	16 841	0.74 ± 0.14	[0.54; 1.07]	13 044

[a] See text for description.
[b] Posterior mean ± standard deviation.
[c] 95% equal-tailed posterior probability interval based on the 2.5th and 97.5th percentiles of the posterior density.
[d] Total effective number of independent samples across the three MCMC chains using Sorensen *et al.* [33].

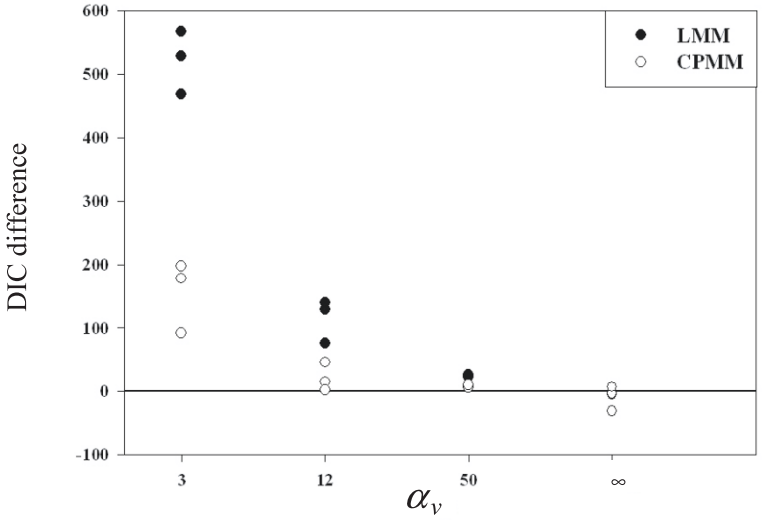

Figure 1. DIC differences between homoskedastic and heteroskedastic error models for linear mixed model (LMM) analysis of normally distributed data and cumulative probit mixed model (CPMM) analysis of ordinal data for 3 replicated datasets from each of four different populations defined by high (α_ν = 3), moderate (α_ν = 12), low (α_ν = 50) or non-existing (α_ν = ∞) levels of random variability of residual heteroskedasticity.

The simulation study was also used to validate DIC as model choice criteria. For this purpose, DIC values based on homoskedastic and heteroskedastic GLMM analyses for each replicated dataset within each of the four populations were determined for the LMM analysis of **L** and for the CPMM analysis of **y**. A DIC difference exceeding 7 has been suggested by Spiegelhalter *et al.* [38] as indication of a decisive difference in model fit. The DIC differences between homoskedastic and heteroskedastic GLMM (LMM for **y** and CPMM for **L**) as matched for each replicate dataset within populations are shown in Figure 1. In all cases, DIC differences were clearly in favor the correct model except when $\alpha_\nu = \infty$ where model choice was always indecisive except for one CPMM case where the homoskedastic error model was correctly chosen. Furthermore, as expected, the magnitude of DIC differences involving LMM analyses of **L** was consistently higher than that for CPMM analyses of **y** as there is likely greater statistical power for inference on continuous **L** as opposed to discrete **y**. Note that the DIC differences between homoskedastic and heteroskedastic error GLMM approached 0 with increasing α_ν, such that there was understandably lower power to conclude the existence of random sources of residual heteroskedasticity when it was very mild.

5.2. Application to calving ease scores in Italian Piemontese cattle

5.2.1. Genetic parameter inference

Sire and maternal grandsire LMM and CPMM were used for the analyses of BW and CE, respectively, from Italian Piemontese cattle. Posterior means and standard deviations and 95% PPI on dispersion parameters, *i.e.* variance components and heteroskedastic parameters, are summarized in Table II, whereas similar quantities are presented for genetic parameters, *i.e.*, heritabilities and their key differences and the direct-maternal genetic correlation, in Table III. The effective sample size (ESS) as a measure of the number of effectively independent samples amongst the 100 000 correlated MCMC samples saved from each chain are also reported for each parameter using Sorensen *et al*. [35]. All reported results in Tables II and III are based on combining information from the three independent MCMC chains after burn-in. All posterior densities (not shown) were nearly symmetric and unimodal which is further illustrated by the near symmetricity of the 95% PPI about the reported posterior means in Tables II and III.

The total number of ESS for dispersion parameters across the three chains ranged from 1836 to 21 839 for the LMM analysis of BW and from 841 to 14 045 for the CPMM analysis of CE scores. As anticipated from the results of the simulation study, the CPMM analysis of CE generated lower ESS than the LMM analysis of BW. The additional inference on α_ν, $\sigma^2_{e_{male}}$, and $\sigma^2_{e_{female}}$ relative to a homoskedastic GLMM specification did not appear to adversely impact MCMC mixing on other dispersion parameters; for example, the reported ESS for the heterogeneous CPMM in Tables II and III appeared to be comparable to those for the homogeneous CPMM on the same data and for the same number of MCMC cycles as reported in Kizilkaya *et al.* [20].

We observed that the posterior means of the residual variance using the homoskedastic error models (results not reported) were nearly equal to the average of the posterior means for $\sigma^2_{e_{male}}$ and $\sigma^2_{e_{female}}$ in the heteroskedastic error models for both GLMM reported in Table II. Note that for BW and CE, the posterior mean of $\sigma^2_{e_{male}}$ was over 40% greater than that for $\sigma^2_{e_{female}}$, and in either case the 95% PPI on $\sigma^2_{e_{male}} - \sigma^2_{e_{female}}$ did not include 0. These significant differences further translated into sex-specific differences for direct and maternal heritabilities estimated to be, respectively, 0.07 and 0.03 greater for BW in females and 0.13 and 0.03 greater for CE in females; furthermore the 95% PPI on each of these differences did not include 0 as also indicated in Table III. Heteroskedastic LMM and CPMM led to relatively large posterior means ($\geq$0.60) and 2.5th percentiles ($\geq$0.46) for $\sigma_\nu = \frac{1}{\sqrt{\alpha_\nu - 2}}$, thereby providing convincing

Table III. Posterior inference on genetic parameters allowing for residual heteroskedasticity based on a linear mixed model analysis of birth weight and a cumulative probit mixed model analysis of calving ease scores in Italian Piemontese cattle.

	Linear mixed model analysis			Cumulative probit mixed model analysis		
Parameter[a]	PM± SD[b]	95% PPI[c]	ESS[d]	PM± SD[b]	95% PPI[c]	ESS[d]
$h^2_{d_male}$	0.25±0.04	[0.17; 0.33]	3392	0.37±0.06	[0.26; 0.50]	2687
$h^2_{m_male}$	0.09±0.03	[0.05; 0.15]	1617	0.09±0.03	[0.04; 0.17]	943
$h^2_{d_female}$	0.32±0.05	[0.22; 0.43]	3315	0.50±0.08	[0.35; 0.68]	2488
$h^2_{m_female}$	0.12±0.04	[0.07; 0.20]	1610	0.13±0.05	[0.05; 0.23]	935
r_{dm}	–0.30±0.15	[–0.57; 0.02]	2328	–0.63±0.15	[–0.89; –0.21]	975
$h^2_{d_female} - h^2_{d_male}$	0.07 ±0.01	[0.05; 0.10]	4354	0.13±0.03	[0.09; 0.19]	3093
$h^2_{m_female} - h^2_{m_male}$	0.03 ±0.01	[0.02; 0.05]	1819	0.03±0.01	[0.01; 0.06]	1025

[a] See text for description.
[b] Posterior mean ± standard deviation.
[c] 95% equal-tailed posterior probability interval based on the 2.5th and 97.5th percentiles of the posterior density.
[d] Total effective number of independent samples across the three MCMC chains using Sorensen *et al.* [33].

evidence that the degree of residual heteroskedasticity across herds is great for both BW and CE.

Approximate 95% PPI (based on the posterior mean ±2 posterior standard deviations) for herd-specific ν_j in the LMM and CPMM analyses of BW and CE, respectively, are presented in Figure 2. For both traits, it can be seen that the approximate 95% PPI for ν_j in many herds do not overlap with the prior expected value of 1, indicating that residual variances for these herds are significantly higher or lower than average, as would be anticipated from the inference on σ_ν in Table II. It should be noted that the ranges of posterior means of ν_j were greater than 10 fold for the LMM analysis of BW in Figure 2a and greater than 20 fold for the CPMM analysis of CE in Figure 2b; *i.e.* residual variances in some herds are estimated to be as large as 20 times greater than residual variances in other herds.

For each homoskedastic and heteroskedastic GLMM, DIC values were averaged across the three independent MCMC chains. Differences in DIC values within models never ranged by more than 3 units across these three chains, thereby indicating satisfactory control of Monte Carlo error. For the LMM analysis of BW, the average DIC value for the heteroskedastic error model was 44 332 compared to 45 240 for the homoskedastic error model, thereby clearly establishing the heteroskedastic error LMM as the better fitting model. Similarly, for CE, the average DIC value for the heteroskedastic CPMM was 15 949 compared to 16 563 for the homoskedastic CPMM, again favoring the heteroskedastic error model. Kizilkaya *et al.* [20] analyzed the same CE data, comparing the homoskedastic error CTMM to the homoskedastic error CPMM and determined that the average DIC value (16 348) for the homoskedastic CTMM made that the model of choice. Using the results from Kizilkaya *et al.* [20] and the results reported here, the heteroskedastic error CPMM appears to be clearly superior to both the homoskedastic error CTMM and CPMM models for fit to CE in Italian Piemontese cattle.

5.2.2. Inferences on sire effects

Posterior means of elements of **s** were used as corresponding point estimates of expected progeny differences (EPD) under both homoskedastic and heteroskedastic error GLMM. Scatterplots of the heteroskedastic error EPD *versus* homoskedastic error EPD for the LMM analysis of BW and for the CPMM analysis of CE are provided in Figure 3. Although the degree of overall reranking was not great between the two error specifications (rank correlations > 0.96 for both GLMM), the same was not necessarily true for the

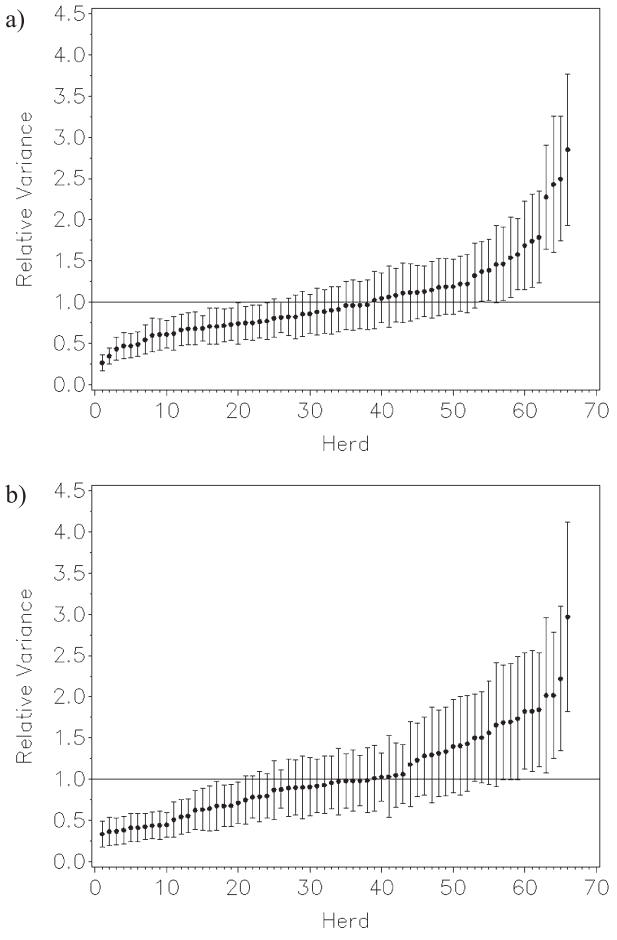

Figure 2. Posterior means (•) and approximate 95% posterior probability interval (bar endpoints) on herd-specific relative variances based on (a) heteroskedastic linear mixed model analysis of birth weights and (b) heteroskedastic cumulative probit mixed model analysis of calving ease scores for each of 66 Italian Piemontese herds labeled by posterior mean rankings within each of the two analyses. Horizontal reference line of a unitary relative variance superimposed.

bottom and top 10% of sires as ranked by the appropriate heteroskedastic error GLMM. For the bottom 10%, rank correlations were less than 0.86 for both GLMM whereas the same correlations were less than 0.80 for the top 10%. EPD differences between the heteroskedastic and homoskedastic error GLMM ranged from –0.74 kg to 0.96 kg for BW and from –0.20 to 0.22 (on the latent scale) for CE; these ranges are not trivial relative to σ_s as based on the

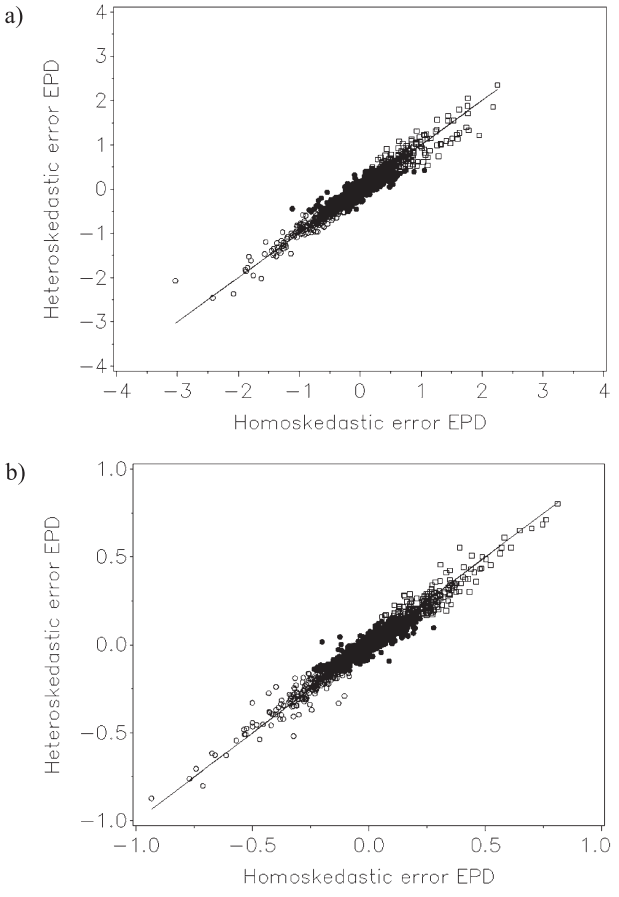

Figure 3. Expected progeny differences (EPD) for heteroskedastic error *versus* homoskedastic error models based on (a) linear mixed model analysis of birthweights and (b) cumulative probit mixed model analysis of calving ease scores in Italian Piemontese cattle for bottom 10% (○), middle 80% (●) and top 10% (□) by heteroskedastic error EPD rankings within each of the two analyses. Reference line of unitary slope and zero intercept superimposed.

square root of the corresponding posterior means of σ_s^2 in Table III. It is interesting to note, in particular, the large EPD discrepancy for BW between the heteroskedastic and homoskedastic error LMM for the left-most point in Figure 3a. This sire had all of his progeny in year-seasons deriving from one herd; incidentally, this herd corresponds to the right-most 95% PPI provided in Figure 2a and having the largest posterior mean for relative residual variance (ν_j).

6. DISCUSSION

Accounting for heterogeneous residual variances in genetic evaluations has represented an important area of research in animal breeding (*e.g.* [7, 9, 11, 17, 18, 26, 29, 31, 32]). We have developed a general framework for structural modeling of heterogeneous residual variances in GLMM based on fully Bayesian inference using MCMC methods. Our method can be readily extended to allow for heterogeneous residual variances across subclasses for other GLMM analyses, particularly those that involve normally distributed latent variables, including censored data models [36].

We validated two different heteroskedastic error GLMM by simulation, demonstrating that fully Bayesian inference allows satisfactory inference on the parameters that specify residual heteroskedasticity. Our simulation study further indicated that Bayesian model choice criteria such as DIC can be used with confidence to choose between heteroskedastic and homoskedastic GLMM error models.

Residual variances for BW and CE were estimated to be over 40% greater in male calves compared to female calves in Italian Piemontese cattle, contributing therefore to smaller heritabilities of direct and maternal effects for male calves. These results for BW are in relatively good agreement with those from Garrick *et al.* [11]. Our results for CE are in agreement with analyses on Blonde d'Aquitaine cattle [8] but are substantially more deviant than those results reported by Ducrocq [7] who estimated residual variances to be 7–18% greater in male calves in French Holstein, Normande and Montbéliarde populations and those results reported by Canavesi *et al.* [3] who estimated residual variances to be less than 3% greater for male calves in Italian Holsteins. Because of the large estimated differences in residual variances, direct and maternal heritabilities were then determined to be substantially larger for female calves than for male calves for both BW and CE. However, our results and particularly its CE comparisons with Ducrocq [7] and Canavesi *et al.* [3] should be treated with caution. Firstly, our population involves a double muscled breed with a high frequency of unfavorable CE scores relative to the cattle populations considered in Ducrocq [7] and Canavesi *et al.* [3]. This distinction appears to be further apparent in that our reported direct heritability estimates for either calf sex were substantially larger than corresponding estimates from recent studies using CPMM [3, 7, 22, 24, 25, 27, 28, 42]. Nevertheless, our inference on a strongly negative direct-maternal genetic correlation for both BW and CE is in good agreement with previous work (*e.g.* [2, 42]). Secondly, Ducrocq [7] and Canavesi *et al.* [3] also modeled constant genetic/residual variance ratios, following the work of Foulley and Gianola [10], whereas we

model constant genetic variances across environments; a more formal comparison using, *e.g.* DIC, of these two specifications is worthy of further study. To our knowledge, nobody has yet inferred upon sex-specific genetic variance for CE. This is an area for further research and requiring a substantially larger dataset than what we considered as the power for inferring upon sex-specific genetic variance is likely to be substantially lower than that for sex-specific residual variance using the CPMM. Finally, Ducrocq [7] and Canavesi *et al.* [3] used approximate marginal maximum likelihood (MML) inference which has been demonstrated to lead to biased estimates of dispersion parameters in simulation studies [15] although our recent work does not suggest meaningful differences between MML and MCMC inference using a CPMM for CE based on larger datasets [19].

There appeared to be no appreciable differences in overall sire rankings based on EPD rank correlations between homoskedastic and heteroskedastic error GLMM for either BW or CE; however, this was not necessarily the case for sires with more extreme EPD for direct genetic effects, specifically the top and bottom 10%. These results are significant in that breeder selection strategies for both traits invariably involve either direct selection of sires in one of these tails and/or avoidance of sires in the other tail, particularly for first parity dams. Furthermore, absolute differences in EPD between the two GLMM for each trait were large relative to the genetic variability. Even greater differences in EPD rankings and absolute differences might be anticipated for genetic evaluations of dams using animal model GLMM since they are used less extensively across potentially heteroskedastic subclasses, *e.g.* herds, compared to sires.

Kizilkaya *et al.* [20] used homoskedastic error CTMM whereas Ducrocq [7], Foulley [8], and Canavesi *et al.* [3] used heteroskedastic error CPMM, all as independent efforts to provide stable genetic evaluations of sires for subjectively recorded CE. Elements of both sets of models, however, may be required. That is, CE data quality may be compromised if, for example, some breeders assign CE scores on a much wider ordinal scale than others, thereby requiring a heteroskedastic error model, or decisions on CE assistance, *e.g.* Caesarean sections, are more likely for highly valued dams, thereby requiring a robust specification like the CTMM. It is conceptually easy to jointly extend the work presented here for heteroskedastic error GLMM and for t-error GLMM as in Kizilkaya *et al.* [20] to develop heteroskedastic t-error LMM for BW or heteroskedastic error CTMM for CE. We have already pursued this work with DIC values clearly pointing to improved model fit based on these enhanced specifications for both BW and CE; this work will be reported in a future study.

The existence of potentially sizable residual and genetic correlations between BW and CE also facilitate more accurate genetic evaluations of sires for CE based on a bivariate threshold/linear multiple trait analysis with BW [23, 42]. We believe the model that we have presented will be useful in facilitating a mixed effects extension to bivariate heteroskedastic error modeling for both BW and CE jointly using, say, inverted Wishart rather than inverted gamma structural prior specifications.

ACKNOWLEDGEMENTS

This research was supported by the Michigan Agricultural Experiment Station (Project MICL 01822). ANABORAPI (Associazione nazionale allevatori bovini di razza Piemontese, Strada Trinità 32a, 12061 Carrù, Italy) is gratefully acknowledged for providing the data for this study. Dr. Ignacy Misztal is acknowledged for his provision of FSPAK90 Fortran subroutines used in this study.

REFERENCES

[1] Albert J., Criticism of a hierarchical Bayes model using Bayes factors, Stat. Med. 18 (1999) 287–305.

[2] Bennett G.L., Gregory K.E., Genetic (co)variances for calving difficulty score in composite and parental populations of beef cattle: I. Calving difficulty score, birth weight, weaning weight, and postweaning gain, J. Anim. Sci. 79 (2001) 45–51.

[3] Canavesi F., Biffani S., Samore A.B., Revising the genetic evaluation for calving ease in the Italian Holstein Friesian. Interbull Bull. 30 (2003) 82–85 http://www-interbull.slu.se/bulletins/framesida-pub.htm.

[4] Chib S., Greenberg E., Understanding the Metropolis Hastings algorithm, Am. Stat. 49 (1995) 327–335.

[5] Clayton D.G., Generalized linear mixed models, in: Gilks W.R., Richardson S., Spiegelhalter D.J. (Eds.), Markov chain Monte Carlo in Practice, Chapman and Hall, London, 1996, pp. 275–302.

[6] Cowles M.K., Accelerating Monte Carlo Markov chain convergence for cumulative link generalized linear models, Stat. Comp. 6 (1996) 101–111.

[7] Ducrocq V., Calving ease evaluation of French dairy bulls with a heteroskedastic threshold model with direct and maternal effects, Interbull Bull. 30 (2000) 82–85 http://www-interbull.slu.se/bulletins/framesida-pub.htm.

[8] Foulley J.L., Heteroskedastic threshold models with applications to the analysis of calving difficulties, Proceedings of the International Workshop on Genetic Improvement of Functional Traits in Cattle: Fertility and Reproduction, Grub Germany, November 1997, International Bull Evaluation Service, Bull. 18 (1998) pp. 3–11.

[9] Foulley J.L., San Cristobal M., Gianola D., Im S., Marginal likelihood and Bayesian approaches to the analysis of heterogeneous residual variances in mixed linear Gaussian models, Comput. Stat. Data. Anal. 13 (1992) 291–305.
[10] Foulley J.L., Gianola D., Statistical analysis of ordered categorical data via a structural heteroskedastic threshold model, Genet. Sel. Evol. 28 (1996) 249–273.
[11] Garrick D.J., Pollak E.J., Quaas R.L., Van Vleck L.D., Variance heterogeneity in direct and maternal weight traits by sex and percent purebred for Simmental-sired calves, J Anim. Sci. 67 (1989) 2515–2528.
[12] Gianola D., Foulley J.L., Sire evaluation for ordered categorical data with a threshold model, Genet. Sel. Evol. 15 (1983) 201–224.
[13] 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 (1991) 2805–2823.
[14] Henderson C.R., Inverse of a matrix of relationships due to sires and maternal grandsires in an inbred population, J. Dairy Sci. 59 (1976) 1585–1588.
[15] Hoeschele I., Tier B., Estimation of variance components of threshold characters by marginal posterior mode and means via Gibbs sampling, Genet. Sel. Evol. 27 (1995) 519–540.
[16] Hill W.G., On selection among groups with heterogeneous variance, Anim. Prod. 39 (1984) 473–477.
[17] Ibanez M.A., Carabano M.J., Alenda R., Identification of sources of heterogeneous residual and genetic variances in milk yield data from the Spanish Holstein-Friesian population and impact on genetic evaluation, Livest. Prod. Sci. 59 (1999) 33–49.
[18] Jaffrezic F., Robert-Granié C., Foulley J.L., A quasi-score approach to the analysis of ordered categorical data via a mixed heteroskedastic threshold model, Genet. Sel. Evol. 31 (1999) 301–318.
[19] Kizilkaya K., Banks B.D., Carnier P., Albera A., Bittante G., Tempelman R.J., Bayesian inference strategies for the prediction of genetic merit using threshold models with an application to calving ease scores in Italian Piemontese cattle, J. Anim. Breed. Genet. 119 (2002) 209–220.
[20] Kizilkaya K., Carnier P., Albera A., Bittante G., Tempelman R.J., An assessment of cumulative *t*-link threshold models for the genetic analysis of calving ease scores, Genet. Sel. Evol. 35 (2003) 489–512.
[21] Littell R.C., Freund R.J., Spector P.J., SAS® System for Linear Models. SAS® Institute Inc., Cary, NC, USA, 1991.
[22] Luo M.F., Boettcher P.J., Dekkers J.C.M., Schaeffer L.R., Bayesian analysis for estimation of genetic parameters of calving ease and stillbirth for Canadian Holsteins, J. Dairy Sci. 82 (1999) 1848.
[23] Luo M.F., Boettcher P.J., Schaeffer L.R., Dekkers J.C.M., Bayesian inference for categorical traits with an application to variance components estimation, J. Dairy Sci. 84 (2001) 694–704.
[24] Manfredi E.J., San Cristobal M., Foulley J.L., Some factor affecting the estimation of genetic parameters for cattle dystocia under a threshold model, Anim. Prod. 53 (1991) 151–156.

[25] Manfredi E.J., Ducrocq V., Foulley J.L., Genetic analysis of dystocia in dairy cattle, J. Dairy Sci. 74 (1991) 1715–1723.
[26] Meuwissen T.H.E., deJong G., Engel B., Joint estimation of breeding values and heterogeneous variances of large data files, J. Dairy Sci. 79 (1996) 310–316.
[27] McGuirk B.J., Going I., Gilmour A.R., The genetic evaluation of beef sires used for crossing with dairy cows in the UK. 2. Genetic parameters and sire merit predictions for calving survey traits, Animal Sci. 66 (1998) 47–54.
[28] McGuirk B.J., Going I., Gilmour A.R., The genetic evaluation of UK Holstein Friesian sires for calving ease and related traits, Animal Sci. 68 (1999) 413–422.
[29] Robert-Granié C., Ducrocq V., Foulley J.L., Heterogeneity of variance for type traits in the Montbeliarde cattle breed, Genet. Sel. Evol. 29 (1997) 545–570.
[30] Robinson G.K., That BLUP is a good thing - the estimation of random effects, Stat. Sci. 6 (1991) 15–51.
[31] San Cristobal M., Foulley J.L., Manfredi E., Inference about multiplicative heteroskedastic components of variance in a mixed linear Gaussian model with an application to beef cattle breeding, Genet. Sel. Evol. 25 (1993) 3–30.
[32] San Cristobal-Gaudy M., Bodin L., Elsen J.-M., Chevalet C., Genetic components of litter size variability in sheep, Genet. Sel. Evol. 33 (2001) 249–271.
[33] Searle S.R., Matrix Algebra Useful for Statistics, John Wiley & Sons, NewYork, 1982.
[34] Sorensen D.A., Gianola D., Likelihood, Bayesian, and MCMC methods in quantitative genetics, Springer-Verlag, New York, 2002.
[35] Sorensen D.A., Andersen S., Gianola D., Korsgaard I., Bayesian inference in threshold models using Gibbs sampling, Genet. Sel. Evol. 27 (1995) 229–249.
[36] Sorensen D.A., Gianola D., Korsgaard I., Bayesian mixed-effects model analysis of a censored data distribution with animal breeding applications, Acta Agric. Scand. Sec. A. 48 (1998) 222–229.
[37] Sorensen D.A., Waagepetersen R., Normal linear models with genetically structured residual heterogeneity: a case study, Genet. Res. Camb. 82 (2003) 207–222.
[38] Spiegelhalter D.J., Best N.G., Carlin B.P., van der Linde A., Bayesian measures of model complexity and fit, J.R. Statist. Soc. B 64 (2002) 1–34.
[39] Tempelman R.J., Generalized linear mixed models in dairy cattle breeding, J. Dairy Sci. 81 (1998) 1428–1444.
[40] Tempelman R.J., Gianola D., A mixed effects model for overdispersed count data in animal breeding, Biometrics 52 (1996) 265–279.
[41] Tempelman R.J., Gianola D., Genetic analysis of fertility in dairy cattle using negative binomial mixed models, J. Dairy Sci. 82 (1999) 1834–1847.
[42] Varona L., Misztal I., Bertrand J.K., Threshold-linear versus linear-linear analysis of birth weight and calving ease using an animal model: I. Variance component estimation, J. Anim. Sci. 77 (1999) 1994–2002.
[43] 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.