Bayesian analysis of genetic change due to selection using Gibbs sampling

Da Sorensen, Cs Wang, J Jensen, D Gianola

▶ **To cite this version:**

Da Sorensen, Cs Wang, J Jensen, D Gianola. Bayesian analysis of genetic change due to selection using Gibbs sampling. Genetics Selection Evolution, 1994, 26 (4), pp.333-360. hal-00894037

HAL Id: hal-00894037
https://hal.science/hal-00894037

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Genet Sel Evol (1994) 26, 333-360
© Elsevier/INRA

Original article

Bayesian analysis of genetic change due to selection using Gibbs sampling

DA Sorensen [1], CS Wang [2], J Jensen [1], D Gianola [2]

[1] *National Institute of Animal Science, Research Center Foulum, PB 39, DK8830 Tjele, Denmark;*
[2] *Department of Meat and Animal Science, University of Wisconsin-Madison, Madison, WI 53706-1284, USA*

(Received 7 June 1993; accepted 10 February 1994)

Summary – A method of analysing response to selection using a Bayesian perspective is presented. The following measures of response to selection were analysed: 1) total response in terms of the difference in additive genetic means between last and first generations; 2) the slope (through the origin) of the regression of mean additive genetic value on generation; 3) the linear regression slope of mean additive genetic value on generation. Inferences are based on marginal posterior distributions of the above-defined measures of genetic response, and uncertainties about fixed effects and variance components are taken into account. The marginal posterior distributions were estimated using the Gibbs sampler. Two simulated data sets with heritability levels 0.2 and 0.5 having 5 cycles of selection were used to illustrate the method. Two analyses were carried out for each data set, with partial data (generations 0–2) and with the whole data. The Bayesian analysis differed from a traditional analysis based on best linear unbiased predictors (BLUP) with an animal model, when the amount of information in the data was small. Inferences about selection response were similar with both methods at high heritability values and using all the data for the analysis. The Bayesian approach correctly assessed the degree of uncertainty associated with insufficient information in the data. A Bayesian analysis using 2 different sets of prior distributions for the variance components showed that inferences differed only when the relative amount of information contributed by the data was small.

response to selection / Bayesian analysis / Gibbs sampling / animal model

Résumé – Analyse bayésienne de la réponse génétique à la sélection à l'aide de l'échantillonnage de Gibbs. *Cet article présente une méthode d'analyse des expériences de sélection dans une perspective bayésienne. Les mesures suivantes des réponses à la sélection ont été analysées: i) la réponse totale, soit la différence des valeurs génétiques additives moyennes entre la dernière et la première génération; ii) la pente (passant par l'origine) de la régression de la valeur génétique additive moyenne en fonction de la génération; iii) la pente de la régression linéaire de la valeur génétique additive moyenne en fonction de la génération. Les inférences sont basées sur les distributions marginales* a posteriori *des mesures de la réponse génétique définies ci-dessus, avec prise en compte des*

incertitudes sur les effets fixés et les composantes de variance. Les distributions marginales a posteriori *ont été estimées à l'aide de l'échantillonnage de Gibbs. Deux ensembles de données simulées avec des héritabilités de 0,2 et 0,5 et 5 cycles de sélection ont été utilisés pour illustrer la méthode. Deux analyses ont été faites sur chaque ensemble de données, avec des données incomplètes (génération 0-2) et avec les données complètes. L'analyse bayésienne différait de l'analyse traditionnelle, basée sur le BLUP avec un modèle animal, quand la quantité d'information utilisée était réduite. Les inférences sur la réponse à la sélection étaient similaires avec les 2 méthodes quand l'héritabilité était élevée et que toutes les données étaient prises en compte dans l'analyse. L'approche bayésienne évaluait correctement le degré d'incertitude lié à une information insuffisante dans les données. Une analyse bayésienne avec 2 distributions* a priori *des composantes de variance a montré que les inférences ne différaient que si la part d'information fournie par les données était faible.*

réponse à la sélection / analyse bayésienne / échantillonnage de Gibbs / modèle animal

INTRODUCTION

Many selection programs in farm animals rely on best linear unbiased predictors (BLUP) using Henderson's (1973) mixed-model equations as a computing device, in order to predict breeding values and rank candidates for selection. With the increasing computing power available and with the development of efficient algorithms for writing the inverse of the additive relationship matrix (Henderson, 1976; Quaas, 1976), 'animal' models have been gradually replacing the originally used sire models. The appeal of 'animal' models is that, given the model, use is made of the information provided by all known additive genetic relationships among individuals. This is important to obtain more precise predictors and to account for the effects of certain forms of selection on prediction and estimation of genetic parameters.

A natural application of 'animal' models has been prediction of the genetic means of cohorts, for example, groups of individuals born in a given time interval such as a year or a generation. These predicted genetic means are typically computed as the average of the BLUP of the genetic values of the appropriate individuals. From these, genetic change can be expressed as, for example, the regression of the mean predicted additive genetic value on time or on appropriate cumulative selection differentials (Blair and Pollak, 1984). In common with selection index, it is assumed in BLUP that the variances of the random effects or ratios thereof are known, so the predictions of breeding values and genetic means depend on such ratios. This, in turn, causes a dependency of the estimators of genetic change derived from 'animal' models on the ratios of the variances of the random effects used as 'priors' for solving the mixed-model equations. This point was first noted by Thompson (1986), who showed in simple settings, that an estimator of realized heritability given by the ratio between the BLUP of total response and the total selection differential leads to estimates that are highly dependent on the value of heritability used as 'prior' in the BLUP analysis. In view of this, it is reasonable

to expect that the statistical properties of the BLUP estimator of response will depend on the method with which the 'prior' heritability is estimated.

In the absence of selection, Kackar and Harville (1981) showed that when the estimators of variance in the mixed-model equations are obtained with even estimators that are translation invariant and functions of the data, unbiased predictors of the breeding value are obtained. No other properties are known and these would be difficult to derive because of the nonlinearity of the predictor. In selected populations, frequentist properties of predictors of breeding value based on estimated variances have not been derived analytically using classical statistical theory. Further, there are no results from classical theory indicating which estimator of heritability ought to be used, even though the restricted maximum likelihood (REML) estimator is an intuitively appealing candidate. This is so because the likelihood function is the same with or without selection, provided that certain conditions are met (Gianola *et al*, 1989; Im *et al*, 1989; Fernando and Gianola, 1990). However, frequentist properties of likelihood-based methods under selection have not been unambiguously characterized. For example, it is not known whether the maximum likelihood estimator is always consistent under selection. Some properties of BLUP-like estimators of response computed by replacing unknown variances by likelihood-type estimates were examined by Sorensen and Kennedy (1986) using computer simulation, and the methodology has been applied recently to analyse designed selection experiments (Meyer and Hill, 1991). Unfortunately, sampling distributions of estimators of response are difficult to derive analytically and one must resort to approximate results, whose validity is difficult to assess. In summary, the problem of exact inferences about genetic change when variances are unknown has not been solved *via* classical statistical methods.

However, this problem has a conceptually simple solution when framed in a Bayesian setting, as suggested by Sorensen and Johansson (1992), drawing from results in Gianola *et al* (1986) and Fernando and Gianola (1990). The starting point is that if the history of the selection process is contained in the data employed in the analysis, then the posterior distribution has the same mathematical form with or without selection (Gianola and Fernando, 1986). Inferences about breeding values (or functions thereof, such as selection response) are made using the marginal posterior distribution of the vector of breeding values or from the marginal posterior distribution of selection response. All other unknown parameters, such as 'fixed effects' and variance components or heritability, are viewed as nuisance parameters and must be integrated out of the joint posterior distribution (Gianola *et al*, 1986). The mean of the posterior distribution of additive genetic values can be viewed as a weighted average of BLUP predictions where the weighting function is the marginal posterior density of heritability.

Estimating selection response by giving all weight to an REML estimate of heritability has been given theoretical justification by Gianola *et al* (1986). When the information in an experiment about heritability is large enough, the marginal posterior distribution of this parameter should be nearly symmetric; the modal value of the marginal posterior distribution of heritability is then a good approximation to its expected value. In this case, the posterior distribution of selection response can be approximated by replacing the unknown heritability by the mode of its marginal

posterior distribution. However, this approximation may be poor if the experiment has little informational content on heritability.

A full implementation of the Bayesian approach to inferences about selection response relies on having the means of carrying out the necessary integrations of joint posterior densities with respect to the nuisance parameters. A Monte-Carlo procedure to carry out these integrations numerically known as Gibbs sampling is now available (Geman and Geman, 1984; Gelfand and Smith, 1990). The procedure has been implemented in an animal breeding context by Wang *et al* (1993a; 1994a,b) in a study of variance component inferences using simulated sire models, and in analyses of litter size in pigs. Application of the Bayesian approach to the analysis of selection experiments yields the marginal posterior distribution of response to selection, from which inferences about it can be made, irrespective of whether variances are unknown. In this paper, we describe Bayesian inference about selection response using animal models where the marginalizations are achieved by means of Gibbs sampling.

MATERIALS AND METHODS

Gibbs sampling

The Gibbs sampler is a technique for generating random vectors from a joint distribution by successively sampling from conditional distributions of all random variables involved in the model. A component of the above vector is a random sample from the appropriate marginal distribution. This numerical technique was introduced in the context of image processing by Geman and Geman (1984) and since then has received much attention in the recent statistical literature (Gelfand and Smith, 1990; Gelfand *et al*, 1990; Gelfand *et al*, 1992; Casella and George, 1992). To illustrate the procedure, let us suppose there are 3 random variables, X, Y and Z, with joint density $p(x, y, z)$; we are interested in obtaining the marginal distributions of X, Y and Z, with densities $p(x)$, $p(y)$ and $p(z)$, respectively. In many cases, the necessary integrations are difficult or impossible to perform algebraically and the Gibbs sampler provides a means of sampling from such a marginal distribution. Our interest is typically in the marginal distributions and the Gibbs sampler proceeds as follows. Let (x_0, y_0, z_0) represent an arbitrary set of starting values for the 3 random variables of interest. The first sample is:

$$x_1 \sim p(x|Y = y_0, Z = z_0)$$

where $p(x|Y = y_0, Z = z_0)$ is the conditional distribution of X given $Y = y_0$ and $Z = z_0$. The second sample is:

$$y_1 \sim p(y|X = x_1, Z = z_0)$$

where x_1 is updated from the first sampling. The third sample is:

$$z_1 \sim p(z|X = x_1, Y = y_1)$$

where y_1 is the realized value of Y obtained in the second sampling. This constitutes the first iteration of the Gibbs sampling. The process of sampling and updating is

repeated k times, where k is known as the length of the Gibbs sequence. As $k \to \infty$, the points of the kth iteration (x_k, y_k, z_k) constitute 1 sample point from $p(x, y, z)$ when viewed jointly, or from $p(x), p(y)$ and $p(z)$ when viewed marginally. In order to obtain m samples, Gelfand and Smith (1990) suggest generating m independent 'Gibbs sequences' from m arbitrary starting values and using the final value at the kth iterate from each sequence as the sample values. This method of Gibbs sampling is known as a multiple-start sampling scheme, or multiple chains. Alternatively, a single long Gibbs sequence (initialized therefore once only) can be generated and every dth observation is extracted (*eg*, Geyer 1992) with the total number of samples saved being m. Animal breeding applications of the short- and long-chain methods are given by Wang *et al* (1993, 1994a,b).

Having obtained the m samples from the marginal distributions, possibly correlated, features of the marginal distribution of interest can be obtained appealing to the ergodic theorem (Geyer, 1992; Smith and Roberts, 1993):

$$\widehat{u}_m = \frac{1}{m} \sum_{i=1}^{m} g(x_i) \qquad [1]$$

where $x_i (i = 1, \ldots, m)$ are the samples from the marginal distribution of x, u is any feature of the marginal distribution, *eg*, mean, variance or, in general, any feature of a function of x, and $g(\cdot)$ is an appropriate operator. For example, if u is the mean of the marginal distribution, $g(\cdot)$ is the identity operator and a consistent estimator of the variance of the marginal distribution is

$$\frac{1}{m} \sum x_i^2 - \left(\sum \frac{x_i}{m} \right)^2$$

(Geyer, 1992). Note that $\widehat{u}_m$ is a Monte-Carlo estimate of u, and the error of this estimator can be made arbitrarily small by increasing m. Another way of obtaining features of a marginal distribution is first to estimate the density using [1], and then compute summary statistics (features) of that distribution from the estimated density.

There are at least 2 ways to estimate a density from the Gibbs samples. One is to use random samples (x_i) to estimate $p(x)$. We consider the normal kernel estimator (*eg*, Silverman, 1986).

$$\widehat{p}(x) = \frac{1}{mh} \sum_{i=1}^{m} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x - x_i}{h} \right)^2 \right] \qquad [2]$$

where $\widehat{p}(x)$ is the estimated density at x, and h is a fixed constant (called the window width) given by the user. The window width determines the smoothness of the estimated curve.

Another way of estimating a density is based on averaging conditional densities (Gelfand and Smith, 1990). From the following standard result:

$$p(x) = \int \int p(x|y, z) p(y, z) dy dz = E_{y,z}[p(x|y, z)]$$

and given knowledge of the full conditional distribution, an estimate of $p(x)$ is obtained from:

$$\widehat{p}(x) = \frac{1}{m}\sum_{i=1}^{m} p(x|Y = y_i, Z = z_i) \qquad [3]$$

where $y_1, z_1; \ldots, y_m, z_m$ are the realized values of final Y, Z samples from each Gibbs sequence. Each pair y_i, z_i constitutes 1 sample, and there are m such pairs. Notice that the x_i are not used to estimate $p(x)$. Estimation of the density of a function of the original variables is accomplished either by applying [2] directly to the samples of the function, or by applying the theory of transformation of random variables in an appropriate conditional density, and then using [3]. In this case, the Gibbs sampling scheme does not need to be rerun, provided the needed samples are saved.

Model

In the present paper we consider a univariate mixed model with 2 variance components for ease of exposition. Extensions to more general mixed models are given by Wang *et al* (1993b; 1994) and by Jensen *et al* (1994). The model is:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \qquad [4]$$

where $\mathbf{y}$ is the data vector of order n by 1; $\mathbf{X}$ is a known incidence matrix of order n by p; $\mathbf{Z}$ is a known incidence matrix of order n by q; $\mathbf{b}$ is a p by 1 vector of uniquely defined 'fixed effects' (so that $\mathbf{X}$ has full column rank); $\mathbf{a}$ is a q by 1 'random' vector representing individual additive genetic values of animals and $\mathbf{e}$ is a vector of random residuals of order n by 1.

The conditional distribution that pertains to the realization of $\mathbf{y}$ is assumed to be:

$$\mathbf{y}|\mathbf{b}, \mathbf{a}, \sigma_e^2 \sim N(\mathbf{Xb} + \mathbf{Za}, \mathbf{H}\sigma_e^2) \qquad [5]$$

where $\mathbf{H}$ is a known n by n matrix, which here will be assumed to be the identity matrix, and σ_e^2 (a scalar) is the unknown variance of the random residuals.

We assume a genetic model in which genes act additively within and between loci, and that there is effectively an infinite number of loci. Under this infinitesimal model, and assuming further initial Hardy–Weinberg and linkage equilibrium, the distribution of additive genetic values conditional on the additive genetic covariance is multivariate normal:

$$\mathbf{a}|\mathbf{A}, \sigma_a^2 \sim N(\mathbf{0}, \mathbf{A}\sigma_a^2) \qquad [6]$$

In [6], $\mathbf{A}$ is the known q by q matrix of additive genetic relationships among animals, and σ_a^2 (a scalar) is the unknown additive genetic variance in the conceptual base population, before selection took place.

The vector $\mathbf{b}$ will be assumed to have a proper prior uniform distribution:

$$p(\mathbf{b}) \propto \text{ constant}, \qquad \mathbf{b}_{\min} \leqslant \mathbf{b} \leqslant \mathbf{b}_{\max} \qquad [7]$$

where $\mathbf{b}_{\min}$ and $\mathbf{b}_{\max}$ are, respectively, the minimum and maximum values which $\mathbf{b}$ can take, *a priori*. Further, $\mathbf{a}$ and $\mathbf{b}$ are assumed to be independent, *a priori*.

To complete the description of the model, the prior distributions of the variance components need to be specified. In order to study the effect of different priors on inferences about heritability and response to selection, 2 sets of prior distributions will be assumed. Firstly, σ_a^2 and σ_e^2 will be assumed to follow independent proper prior uniform distributions of the form:

$$p(\sigma_i^2) \propto \text{ constant}, \qquad 0 \leqslant \sigma_i^2 < \sigma_{i\max}^2 \qquad (i = e, a) \qquad [8]$$

where $\sigma_{a\max}^2$ and $\sigma_{e\max}^2$ are the maximum values which, according to prior knowledge, σ_a^2 and σ_e^2 can take. In the rest of the paper, all densities which are functions of $\mathbf{b}, \sigma_a^2$, and σ_e^2 will implicity take the value zero if the bounds in [7] and [8] are exceeded.

Secondly, σ_a^2 and σ_e^2 will be assumed to follow *a priori* scaled inverted chi-square distributions:

$$p\left(\sigma_i^2|\nu_i, S_i^2\right) \propto \left(\sigma_i^2\right)^{-((\nu_i/2)+1)} \exp\left[-\frac{\nu_i S_i^2}{2\sigma_i^2}\right] \qquad (i = a, e) \qquad [9]$$

where ν_i and S_i^2 are parameters of the distribution. Note that a uniform prior can be obtained from [9] by setting $\nu_i = -2$ and $S_i^2 = 0$.

Posterior distribution of selection response

Let α represent the vector of parameters associated with the model. Bayes theorem provides a means of deriving the posterior distributions of α conditional on the data:

$$p(\alpha|\mathbf{y}) = \frac{p(\mathbf{y}|\alpha)p(\alpha)}{\int_{R_\alpha} p(\mathbf{y}, \alpha)d\alpha} \qquad [10]$$

The first term in the numerator of the right-hand side of [10] is the conditional density of the data given the parameters, and the second is the prior joint distribution of the parameters in the model. The denominator in [10] is a normalizing constant (marginal distribution of the data) that does not depend on α. Applying [10], and assuming the set of priors [9] for the variance components, the joint posterior distribution of the parameters is:

$$\begin{aligned} p(\mathbf{b}, \mathbf{a}, \sigma_a^2, \sigma_e^2|\mathbf{y}, \nu, \mathbf{S}) &\propto p(\mathbf{y}|\mathbf{b}, \mathbf{a}, \sigma_a^2, \sigma_e^2)p(\mathbf{b}, \mathbf{a}, \sigma_a^2, \sigma_e^2|\nu, \mathbf{S}) \\ &= p(\mathbf{y}|\mathbf{b}, \mathbf{a}, \sigma_a^2, \sigma_e^2)p(\mathbf{b})p(\mathbf{a}|\sigma_a^2)p(\sigma_a^2|\nu_a, S_a^2)p(\sigma_e^2|\nu_e, S_e^2) \end{aligned} \qquad [11]$$

where $\nu = (\nu_a, \nu_e)$ and $\mathbf{S} = \left(S_a^2, S_e^2\right)$. Using [5], [6], [7] and [9] in [11] yields as the joint posterior density:

$$p(\mathbf{b},\mathbf{a},\sigma_a^2,\sigma_e^2|\mathbf{y},\nu,\mathbf{S}) \propto \left(\sigma_e^2\right)^{-\left(\frac{n+\nu_e}{2}+1\right)} \times \exp\left(-\frac{(\mathbf{y}-\mathbf{Xb}-\mathbf{Za})'(\mathbf{y}-\mathbf{Xb}-\mathbf{Za})+\nu_e S_e^2}{2\sigma_e^2}\right) \times \left(\sigma_a^2\right)^{-\left(\frac{q+\nu_a}{2}+1\right)} \exp\left(-\frac{\mathbf{a}'\mathbf{A}^{-1}\mathbf{a}+\nu_a S_a^2}{2\sigma_a^2}\right) \quad [12]$$

The joint posterior under a model assuming the proper set of prior uniform distributions [8] for the variance components is simply obtained by setting $\nu_i = -2$ and $S_i^2 = 0$ $(i = e, a)$ in [12]. To make the notation less burdensome, we drop from now onwards the conditioning on ν and $\mathbf{S}$.

Inferences about response to selection can be made working with a function of the marginal posterior distribution of $\mathbf{a}$. The latter is obtained integrating [12] over the remaining parameters:

$$p(\mathbf{a}|\mathbf{y}) = \int_{R_{b,\Sigma}} \int p(\mathbf{a}|\mathbf{b},\Sigma,\mathbf{y}) p(\mathbf{b},\Sigma|\mathbf{y}) d\mathbf{b}\, d\Sigma = E_{\mathbf{b},\Sigma|\mathbf{y}}[p(\mathbf{a}|\mathbf{b},\Sigma,\mathbf{y})] \quad [13]$$

where $\Sigma = (\sigma_a^2, \sigma_e^2)$ and the expectation is taken over the joint posterior distribution of the vector of 'fixed effects' and variance components. This density cannot be written in a closed form. In finite samples, the posterior distribution of $\mathbf{a}$ should be neither normal nor symmetric.

Response to selection is defined as a linear function of the vector of additive genetic values:

$$\mathbf{R} = \mathbf{Ka} \quad [14]$$

where $\mathbf{K}$ is an appropriately defined transformation matrix and $\mathbf{R}$ can be a vector (or a scalar) whose elements could be the mean additive genetic values of each generation, or contrasts between these means, or alternatively, regression coefficients representing linear and quadratic changes of genetic means with respect to some measure of time, such as generations. By virtue of the central limit theorem, the posterior distribution of $\mathbf{R}$ should be approximately normal, even if [13] is not.

Full conditional posterior distributions (the Gibbs sampler)

In order to implement the Gibbs sampling scheme, the full conditional posterior densities of all the parameters in the model must be obtained. These distributions are, in principle, obtained dividing [12] by the appropriate posterior density function or the equivalent, regarding all parameters in [12] other than the one of interest as known. For the fixed and random effects though, it is easier to proceed as follows.

Define $\mathbf{W}' = (\mathbf{X}\,\mathbf{Z})'; \theta' = (\mathbf{b}'\mathbf{a}'); \Omega = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{-1}k \end{bmatrix}$, where $k = \frac{\sigma_e^2}{\sigma_a^2}$ and write model [4] as:

$$\mathbf{y} = \mathbf{W}\theta + \mathbf{e} \quad [15]$$

Using results from Lindley and Smith (1972), the conditional distribution of θ given all other parameters is multivariate normal:

$$\theta|\boldsymbol{\Sigma}, \mathbf{y} \sim \mathbf{N}(\widehat{\theta}, \mathbf{Q}^{-1}\sigma_e^2) \quad [16]$$

where

$$\mathbf{Q}(\mathbf{W}'\mathbf{W} + \boldsymbol{\Omega}) \quad [17]$$

and $\widehat{\theta}$ satisfies:

$$\mathbf{Q}\widehat{\theta} = \mathbf{W}'\mathbf{y} \quad [18]$$

which are the mixed-model equations of Henderson (1973).

Partitioning $\theta' = (\theta_1'\theta_2'), \mathbf{W}' = (\mathbf{W}_1\mathbf{W}_2)'$ and

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}$$

and using standard multivariate normal theory, it can be shown for any such partition that:

$$\theta_1|\theta_2, \Sigma, \mathbf{y} \sim \mathbf{N}(\widehat{\theta}_1, (\mathbf{Q}_{11})^{-1}\sigma_e^2) \quad [19]$$

where $\widehat{\theta}_1$ is given by:

$$\widehat{\theta}_1 = (\mathbf{Q}_{11})^{-1}(\mathbf{W}_1'\mathbf{y} - \mathbf{Q}_{12}\theta_2) \quad [20]$$

Expressions [19] and [20] can be computed in matrix or scalar form. Let b_i be a scalar corresponding to the ith element in the vector of 'fixed effects', $\mathbf{b}_{-i}$ be $\mathbf{b}$ without its ith element, $\mathbf{x}_i$ be the ith column vector in $\mathbf{X}$, and $\mathbf{X}_{-i}$ be that part of matrix $\mathbf{X}$ with $\mathbf{x}_i$ excluded. Using [19] we find that the full conditional distribution of b_i given all other parameters is normal.

$$b_i|\mathbf{b}_{-i}, \mathbf{a}, \boldsymbol{\Sigma}, \mathbf{y} \sim N\left(\widehat{b}_i, (\mathbf{x}_i'\mathbf{x}_i)^{-1}\sigma_e^2\right) \quad [21]$$

where $\widehat{b}_i$ satisfies:

$$\mathbf{x}_i'\mathbf{x}_i\widehat{b}_i = \mathbf{x}_i'(\mathbf{y} - \mathbf{X}_{-i}\mathbf{b}_{-i} - \mathbf{Z}\mathbf{a}) \quad [22]$$

For the 'random effects', again using [19] and letting $\mathbf{a}_{-i}$ be $\mathbf{a}$ without its ith element, we find that the full conditional distribution of the scalar a_i given all the other parameters is also normal:

$$a_i|\mathbf{b}, \mathbf{a}_{-i}, \boldsymbol{\Sigma}, \mathbf{y} \sim N\left(\widehat{a}_i, (\mathbf{z}_i'\mathbf{z}_i + c_{ii}k)^{-1}\sigma_e^2\right) \quad [23]$$

where $\mathbf{z}_i$ is the ith column of $\mathbf{Z}$, $c_{ii} = (\mathrm{Var}(a_i|\mathbf{a}_{-i}))^{-1}\sigma_a^2$ is the element in the ith row and column of $\mathbf{A}^{-1}$, and $\widehat{a}_i$ satisfies:

$$(\mathbf{z}_i'\mathbf{z}_i + c_{ii}k)\,\widehat{a}_i = \mathbf{z}_i'(\mathbf{y} - \mathbf{X}\mathbf{b}) - k\,\mathbf{c}_{i,-i}\mathbf{a}_{-i} \quad [24]$$

In [24], $\mathbf{c}_{i,-i}$ is the row of $\mathbf{A}^{-1}$ corresponding to the ith individual with the ith element excluded.

The full conditional distribution of each of the variance components is readily obtained by dividing [12] by $p(\mathbf{b}, \mathbf{a}, \boldsymbol{\Sigma}_{-i}|\mathbf{y})$, where $\boldsymbol{\Sigma}_{-i}$ is $\boldsymbol{\Sigma}$ with σ_i^2 excluded. Since this last distribution does not depend on σ_i^2, this becomes (Wang *et al*, 1994a):

$$p\left(\sigma_i^2|\mathbf{b}, \mathbf{a}, \boldsymbol{\Sigma}_{-i}, \mathbf{y}, \nu, \mathbf{S}\right) \propto \left(\sigma_i^2\right)^{-\left(\frac{\widetilde{\nu}_i}{2}+1\right)} \exp\left(-\frac{\widetilde{\nu}_i \widetilde{\mathbf{S}}_i^2}{2\sigma_i^2}\right) \qquad (i = e, a) \qquad [25]$$

which has the form of an inverted chi-square distribution, with parameters:

$$\begin{aligned} \widetilde{\nu}_e &= n + \nu_e;\ \widetilde{\nu}_a = q + \nu_a \\ \widetilde{S}_e^2 &= \frac{(\mathbf{y} - \mathbf{Xb} - \mathbf{Za})'(\mathbf{y} - \mathbf{Xb} - \mathbf{Za}) + \nu_e S_e^2}{\widetilde{\nu}_e} \\ \widetilde{S}_a^2 &= (\mathbf{a}'\mathbf{A}^{-1}\mathbf{a} + \nu_a S_a^2)/\widetilde{\nu}_a \end{aligned} \qquad [26]$$

When the prior distribution for the variance components is assumed to be uniform, the full conditional distributions have the same form as in [25], except that, in [26] and subsequent formulae, $\nu_i = -2$ and $S_i^2 = 0$ $(i = e, a)$.

Generation of random samples from marginal posterior distributions using Gibbs sampling

In this section we describe how random samples can be generated indirectly from the joint distribution [12] by sampling from the conditional distributions [21], [23] and [25]. The Gibbs sampler works as follows:

(i) set arbitrary initial values for $\mathbf{b}, \mathbf{a}$ and Σ;
(ii) sample from [21], and update $b_i, i = 1, \ldots, p$;
(iii) sample from [23] and update $a_i,\ i = 1, \ldots, q$;
(iv) sample from [25] and update σ_a^2;
(v) sample from [25] and update σ_e^2;
(vi) repeat (ii) to (v) k (length of chain) times.

As $k \to \infty$, this creates a Markov chain with an equilibrium distribution having [12] as density. If along the single path k, m samples are extracted at intervals of length d, the algorithm is called a single-long-chain algorithm. If, on the other hand, m independent chains are implemented, each of length k, and the kth iteration is saved as a sample, then this is known as the short-chain algorithm.

If the Gibbs sampler reached convergence, for the m samples $(\mathbf{b}, \mathbf{a}, \boldsymbol{\Sigma})_1, \ldots, (\mathbf{b}, \mathbf{a}, \boldsymbol{\Sigma})_m$ we have:

$$\begin{aligned} &\mathbf{b}_1, \ldots, \mathbf{b}_m \sim p(\mathbf{b}|\mathbf{y}) \\ &\mathbf{a}_1, \ldots, \mathbf{a}_m \sim p(\mathbf{a}|\mathbf{y}) \\ &(\sigma_a^2)_1, \ldots, (\sigma_a^2)_m \sim p(\sigma_a^2|\mathbf{y}) \\ &(\sigma_e^2)_1, \ldots, (\sigma_e^2)_m \sim p(\sigma_e^2|\mathbf{y}) \end{aligned}$$

where $\sim$ means distributed with marginal posterior density $p(\cdot|\mathbf{y})$. In any particular sample, we notice that the elements of the vectors $\mathbf{b}$ and $\mathbf{a}$, b_i and a_i, say, are samples

of the univariate marginal distributions $p(b_i|\mathbf{y})$ and $p(a_i|\mathbf{y})$. In order to estimate marginal posterior densities of the variance components, in addition to the above quantities the following sums of squares must be stored for each of the m samples:

$$(S_e^2)_1, \ldots, (S_e^2)_m$$
$$(S_a^2)_1, \ldots, (S_a^2)_m$$

For estimation of selection response, the quantities to be stored depend on the structure of $\mathbf{K}$ in [14].

Density estimation

As indicated previously, a marginal density can be estimated, for example, using the normal density kernel estimator [2] with the m Gibbs samples from the relevant marginal distribution, or by averaging over conditional densities (equation [3]). Density estimation by the former method is straightforward, by applying [2]. Here, we outline density estimation using [3].

The formulae for variance components and functions thereof were given by Wang *et al* (1993, 1994b). For each of the 2 variance components, the conditional density estimators are:

$$\widehat{p}(\sigma_a^2|\mathbf{y}) = \frac{1}{m}\sum_{j=1}^{m} p(\sigma_a^2|(\mathbf{b})_j, (\mathbf{a})_j, (\sigma_e^2)_j, \mathbf{y})$$

$$= \left[\Gamma\left(\frac{\widetilde{\nu}_a}{2}\right)\right]^{-1}\left[\frac{\widetilde{\nu}_a}{2}\right]^{\widetilde{\nu}_a/2}(\sigma_a^2)^{-[(\widetilde{\nu}_a/2)+1]}\frac{1}{m}\sum_{j=1}^{m}\left[\left(\widetilde{S}_a^2\right)^{\widetilde{\nu}_a/2}\right]_j \exp\left[-\frac{\widetilde{\nu}_a(\widetilde{S}_a^2)_j}{2\sigma_a^2}\right] \quad [27]$$

and for the error variance component:

$$\widehat{p}(\sigma_e^2|\mathbf{y}) = \left[\Gamma\left(\frac{\widetilde{\nu}_e}{2}\right)\right]^{-1}\left[\frac{\widetilde{\nu}_e}{2}\right]^{\widetilde{\nu}_e/2}(\sigma_e^2)^{-[(\widetilde{\nu}_e/2)+1]}$$
$$\frac{1}{m}\sum_{j=1}^{m}\left[\left(\widetilde{S}_e^2\right)^{\widetilde{\nu}_e/2}\right]_j \exp\left[-\frac{\widetilde{\nu}_e\left(\widetilde{S}_e^2\right)_j}{2\sigma_e^2}\right] \quad [28]$$

The estimated values of the density are obtained by fixing σ_a^2 and σ_e^2 at a number of points and then evaluating [27] and [28] at each point over the m samples. Notice that the realized values of σ_a^2 and σ_e^2 obtained in the Gibbs sampler are not used to estimate the marginal posterior densities in [27] and [28].

To estimate the marginal posterior density of heritability $h^2 = \sigma_a^2/(\sigma_a^2+\sigma_e^2)$, the point of departure is the full conditional distribution of σ_a^2. This distribution has σ_e^2 as a conditioning variable, and therefore σ_e^2 is treated as a constant. Since the inverse transformation is $\sigma_a^2 = \sigma_e^2 h^2/(1-h^2)$, and the Jacobian of the transformation from σ_a^2 to h^2 is $J = \sigma_e^2/(1-h^2)^2$, then from [27] we obtain:

$$\widehat{p}(h^2|\mathbf{y}) = \left[\Gamma\left(\frac{\widetilde{\nu}_a}{2}\right)\right]^{-1}\left[\frac{\widetilde{\nu}_a}{2}\right]^{\widetilde{\nu}_a/2}\left(h^2\right)^{-[(\widetilde{\nu}_a/2)+1]}\left(1-h^2\right)^{(\widetilde{\nu}_a/2)-1}$$
$$\frac{1}{m}\sum_{j=1}^{m}\left[\left(\widetilde{S}_a^2\right)^{(\widetilde{\nu}_a/2)}\right]_j\left[\left(\sigma_e^2\right)^{-(\widetilde{\nu}_a/2)}\right]_j\exp\left[-\frac{1-h^2}{2h^2}\widetilde{\nu}_a\left(\widetilde{S}_a^2\right)_j/\left(\sigma_e^2\right)_j\right] \quad [29]$$

The estimation of the marginal posterior density of each additive genetic value follows the same principles:

$$\widehat{p}(a_i|\mathbf{b},\mathbf{a}_{-i},\boldsymbol{\Sigma},\mathbf{y}) = \frac{1}{m}\sum_{j=1}^{m}p(a_i|(\mathbf{b})_j,(\mathbf{a}_{-i})_j,(\Sigma)_j,\mathbf{y})$$
$$= \frac{1}{m}\sum_{j=1}^{m}\left[\frac{1}{(2\pi)\left[\mathbf{z}_i'\mathbf{z}_i+c_{ii}(k)_j\right]^{-1}\left(\sigma_e^2\right)_j}\right]^{1/2}$$
$$\exp\left[-\frac{1}{2\left[\mathbf{z}_i'\mathbf{z}_i+c_{ii}(k)_j\right]^{-1}\left(\sigma_e^2\right)_j}\left(a_i-(\widehat{a}_i)_j\right)^2\right] \quad [30]$$

where, for the jth sample:

$$(\widehat{a}_i)_j = \left[\mathbf{z}_i'\mathbf{z}_i+c_{ii}(k)_j\right]^{-1}\mathbf{z}_i'\left(\mathbf{y}-\mathbf{X}(\mathbf{b})_j\right)-(k)_j\mathbf{c}_{i,-i}(\mathbf{a}_{-i})_j \quad [31]$$

Equally spaced points in the effective range of a_i are chosen, and for each point an estimate of its density is obtained by inserting, for each of the m Gibbs samples, the realized values for $\sigma_e^2, \mathbf{b}, \mathbf{a}_{-i}$, and k in [30] and [31]. These quantities, together with $\widehat{a}_i$, need to be stored in order to obtain an estimate of the marginal posterior density of the ith additive genetic value. This process is repeated for each of the equally spaced points.

Estimation of the marginal posterior density of response depends on the way it is expressed. In general $\mathbf{R}$ in [14] can be a scalar (the genetic mean of a given generation, or response per time unit) or it can be a vector, whose elements could describe, for example, linear and higher order terms of changes of additive genetic values with time or unit of selection pressure applied. We first derive a general expression for estimation of the marginal posterior density of $\mathbf{R}$ and then look at some special cases to illustrate the procedure.

Assume that $\mathbf{R}$ contains s elements and we wish to estimate $p(\mathbf{R}|\mathbf{y})$ as:

$$\widehat{p}(\mathbf{R}|\mathbf{y}) = \frac{1}{m}\sum_{j=1}^{m}p(\mathbf{R}|(\mathbf{b})_j,(\mathbf{a}_{-i})_j,(\boldsymbol{\Sigma})_j,\mathbf{y}) \quad [32]$$

In order to implement [32], the full conditional distribution on the right-hand side is needed. This distribution is obtained applying the theory of transformations to $p(\mathbf{a}_i|\mathbf{b},\mathbf{a}_{-i},\boldsymbol{\Sigma},\mathbf{y})$, where $\mathbf{a}_i$ is a vector of additive genetic values of order s, and $\mathbf{a}_{-i}$ is the vector of all additive genetic values with $\mathbf{a}_i$ deleted, such that $\mathbf{a} = (\mathbf{a}_i,\mathbf{a}_{-i})'$.

The s additive genetic values in $\mathbf{a}_i$ must be chosen so that the matrix of the transformation from $\mathbf{a}_i$ to $\mathbf{R}$ is non-singular. We can then write [14] as:

$$\mathbf{R} = \mathbf{K}_i\mathbf{a}_i + \mathbf{K}_{-i}\mathbf{a}_{-i} \qquad [33]$$

so that we have expressed $\mathbf{R}$ in a part which is a function of $\mathbf{a}_i$ and another one which is not. The matrix $\mathbf{K}_i$ is non-singular of order s by s, and from [33], the inverse transformation is

$$\mathbf{a}_i = (\mathbf{K}_i)^{-1}(\mathbf{R} - \mathbf{K}_{-i}\mathbf{a}_{-i})$$

Since the Jacobian of the transformation from $\mathbf{a}_i$ to $\mathbf{R}$ is $\det(\mathbf{K}_i^{-1})$, letting $\mathbf{V}_i = \mathrm{Var}(\mathbf{a}_i|\mathbf{b}, \mathbf{a}_{-i}, \boldsymbol{\Sigma}, \mathbf{y})$, and using standard theory, we obtain the result that the conditional posterior distribution of response is normal:

$$\mathbf{R}|\mathbf{b}, \mathbf{a}_{-i}, \boldsymbol{\Sigma}, \mathbf{y} \sim N(\mathbf{K}_i\widehat{\mathbf{a}}_i + \mathbf{K}_{-i}\mathbf{a}_{-i}, \mathbf{K}_i\mathbf{V}_i\mathbf{K}_i') \qquad [34]$$

where, using [19] and [20], it can be shown that:

$$(\mathbf{Z}_i'\mathbf{Z}_i + \mathbf{C}_i k)\widehat{\mathbf{a}}_i = \mathbf{Z}_i'(\mathbf{y} - \mathbf{X}\mathbf{b}) - k\mathbf{C}_{i,-i}\mathbf{a}_{-i} \qquad [35]$$

and

$$\mathbf{V}_i = (\mathbf{Z}_i'\mathbf{Z}_i + \mathbf{C}_i k)^{-1}\sigma_e^2 \qquad [36]$$

In [35] and [36], $\mathbf{C}_i$ is the s by s block of the inverse of the additive genetic relationship matrix whose rows and columns correspond to the elements in $\mathbf{a}_i$, and $\mathbf{C}_{i,-i}$ is the s by $(q-s)$ block associating the elements of $\mathbf{a}_i$ with those of $\mathbf{a}_{-i}$. With [34] available, the marginal posterior distribution of $\mathbf{R}$ can be obtained from [32].

As a simple illustration, consider the estimation of the marginal posterior density of total response to selection, defined as the difference in average additive genetic values between the last generation (f) and the first one:

$$R = \frac{1}{n_f}\sum_{j=1}^{n_f} a_{fj} - \frac{1}{n_1}\sum_{j=1}^{n_1} a_{1j} \qquad [37]$$

where a_{fj} and a_{1j} are additive genetic values of individuals in the final and first generations, and n_f and n_1 are the number of additive genetic values in the final and first generation, respectively. We will arbitrarily choose a_{f1} and carry out a linear transformation from $p(a_{f1}|\mathbf{b}, \mathbf{a}_{-f1}, \boldsymbol{\Sigma}, \mathbf{y})$ to $p(R|\mathbf{b}, \mathbf{a}_{-f1}, \boldsymbol{\Sigma}, \mathbf{y})$. We write [37] in the form of [33]:

$$R = \frac{1}{n_f}\left[a_{f1} + \sum_{j=2}^{n_f} a_{fj}\right] - \frac{1}{n_1}\sum_{j=1}^{n_1} a_{1j} \qquad [38]$$

so that in the notation of [33], we have:

$$K_{f1}a_{f1} = \frac{1}{n_f}a_{f1} \qquad [39]$$

and

$$\mathbf{K}_{-f1}\hat{\mathbf{a}}_{-f1} = \frac{1}{n_f}\sum_{j=2}^{n_f} a_{fj} - \frac{1}{n_1}\sum_{j=1}^{n_1} a_{1j} \qquad [40]$$

Now, since $V_{f1} = \text{Var}(a_{f1}|\mathbf{b}, \mathbf{a}_{-f1}, \boldsymbol{\Sigma}, \mathbf{y}) = (\mathbf{z}'_{f1}\mathbf{z}_{f1} + c_{f1}k)^{-1}\sigma_e^2$, and in [34], $K_{f1}V_{f1}K'_{f1} = V_{f1}/n_f^2$, we have:

$$R|\mathbf{b}, \mathbf{a}_{-f1}, \boldsymbol{\Sigma}, \mathbf{y} \sim N\left[\left(\frac{1}{n_f}\sum_{j=2}^{n_f} a_{fj} - \frac{1}{n_1}\sum_{j=1}^{n_1} a_{1j} + \frac{1}{n_f}\hat{a}_{f1}\right), \left(\frac{\sigma_e^2}{n_f^2}(\mathbf{z}'_{f1}\mathbf{z}_{f1} + c_{f1}k)^{-1}\right)\right] \qquad [41]$$

and the marginal posterior density is estimated as follows:

$$\hat{p}(R|\mathbf{y}) = \frac{1}{m}\sum_{j=1}^{m} p(R|(\mathbf{b})_j, (\mathbf{a}_{-f1})_j, (\boldsymbol{\Sigma})_j, \mathbf{y}) \qquad [42]$$

As a second illustration we consider the case where response is expressed as the linear regression (through the origin) of additive genetic values on time units. Thus R is a scalar. We assume as before that there are f time units, and the response expressed in the form of [33] and [38] is:

$$R = \left[\sum_{j=1}^{f} n_j j^2\right]^{-1} n_f a_{f1} + \left[\sum_{j=1}^{f} n_j j^2\right]^{-1}\left[n_f \sum_{j=2}^{n_f} a_{fj} + \sum_{j=1}^{f-1} n_j a_{j\cdot}\right] \qquad [43]$$

where $a_{j\cdot} = \sum_{k=1}^{n_j} a_{jk}$, the sum of the additive genetic values in time unit j, and n_j is their number. A little manipulation shows that:

$$R|\mathbf{b}, \mathbf{a}_{-f1}, \boldsymbol{\Sigma}, \mathbf{y} \sim N\left[\left[\sum_{j=1}^{f} n_j j^2\right]^{-1}\left[n_f \sum_{j=2}^{n_f} a_{fj} + \sum_{j=1}^{f-1} n_j a_{j\cdot} + n_f\hat{a}_{f1}\right], \left[\sum_{j=1}^{f} n_j j^2\right]^{-2} n_f^2(\mathbf{Z}'_{f1}\mathbf{Z}_{f1} + c_{f1}k)^{-1}\sigma_e^2\right] \qquad [44]$$

EXAMPLES

To illustrate, the methodology was used to analyse 2 simulated data sets. Genotypes were sampled using a Gaussian additive genetic model. The phenotypic variance was always 10, and base population heritability was 0.2 and 0.5 in data sets 1 and 2, respectively. A phenotypic record was obtained by summing a fixed effect

(30 fixed effects in the complete data set, but these values are of no importance here), an additive genetic effect and a residual term. Both additive genetic and residual effects were assumed to follow normal distributions with null means and variances based on the 2 heritability levels. In the generation of Mendelian sampling effects, parental inbreeding was taken into account.

Five cycles of single trait selection based on a BLUP animal model were practised. Each generation resulted from the mating of 8 males and 20 females, and each female produced 3 offspring with records and 2 additional female offspring without records which were also considered as female replacements. Males were selected across generations (the best 8 each cycle) and females, which bred only once, were selected from the highest scoring predicted breeding values available each generation. The total number of animals at the end of the program was 968, of which 720 had records. The total number of sires and dams in each data set was 42 and 241. Details of the simulation including a description of the generation of fixed effects and additive genetic values can be found in Sorensen (1988).

Two analyses were carried out for each data set: an analysis based on all records; and one using data from generations 0–2 only. The rationale for this is that it is not uncommon in the literature to find reports of selection experiments in progress. Hence, we can illustrate in this manner how inferences are modified as additional information is collected.

For each analysis, marginal posterior densities for the additive genetic variance (σ_a^2), residual variance (σ_e^2), phenotypic variance (σ_p^2) and heritability (h^2) were estimated. In addition, marginal posterior densities of 2 expressions for selection response were estimated. These were: 1) difference between final and initial average additive genetic values or total response (TR); and 2) linear regression through the origin of the linear regression of additive genetic values on generation. Unless otherwise stated, the analyses reported below were performed assuming that variance components follow *a priori* independent uniform distributions as shown in [8]. The upper bounds of the additive and environmental variances ($\sigma^2_{a\max}, \sigma^2_{e\max}$), were arbitrarily chosen to be 10 and 20, respectively.

The Gibbs sampler was implemented using a single chain of total length 1 205 000, with a warm-up (initial iterations discarded) of length 5 000, and a sub-sampling interval between samples of $d = 10$. Thus, a total of $m = 120\,000$ samples were saved. Calculations were programmed in Fortran in double precision. Random numbers were generated using IMSL subroutines (IMSL Inc, 1989).

Plots of the estimated marginal posterior densities of σ_a^2, σ_e^2, h^2, TR and δ were obtained with both the average of conditionals and with the normal kernel density estimator; those for σ_p^2, δ_0 and δ_1 were generated using the normal kernel density estimator only, for technical reasons. The window used for the normal kernel estimator was the range of the effective domain of the parameter, divided by 75. The effective domain covered 99.9% of the density mass. Each of the plots was generated by dividing the effective domain of that variable into 100 evenly spaced intervals. Summary statistics of distributions, such as the mean, median and variance were computed by Simpson's integration rules, by further dividing the effective domain into 1 000 evenly spaced intervals using cubic-spline techniques (IMSL Inc, 1989). Modes were located through grid search. For illustration and not as a means of a definite comparison between methods, a standard classical analysis

was carried out. First, REML estimates of variance components were obtained. These REML estimates were then used in lieu of the true values in Henderson's mixed-model equations to obtain estimates of additive genetic values. Finally, estimated genetic responses to selection were computed based on appropriate averages of the estimated additive genetic values. Results of this classical analysis are reported later under the heading ST in figures 5–8. The estimated marginal posterior densities of the variance components and of heritability are shown in figures 1 and 2 for the analysis based on partial (PART) and whole (WHOLE) data, respectively, when base population heritability was 0.2. Corresponding plots for base heritability 0.5 are given in figures 3 and 4.

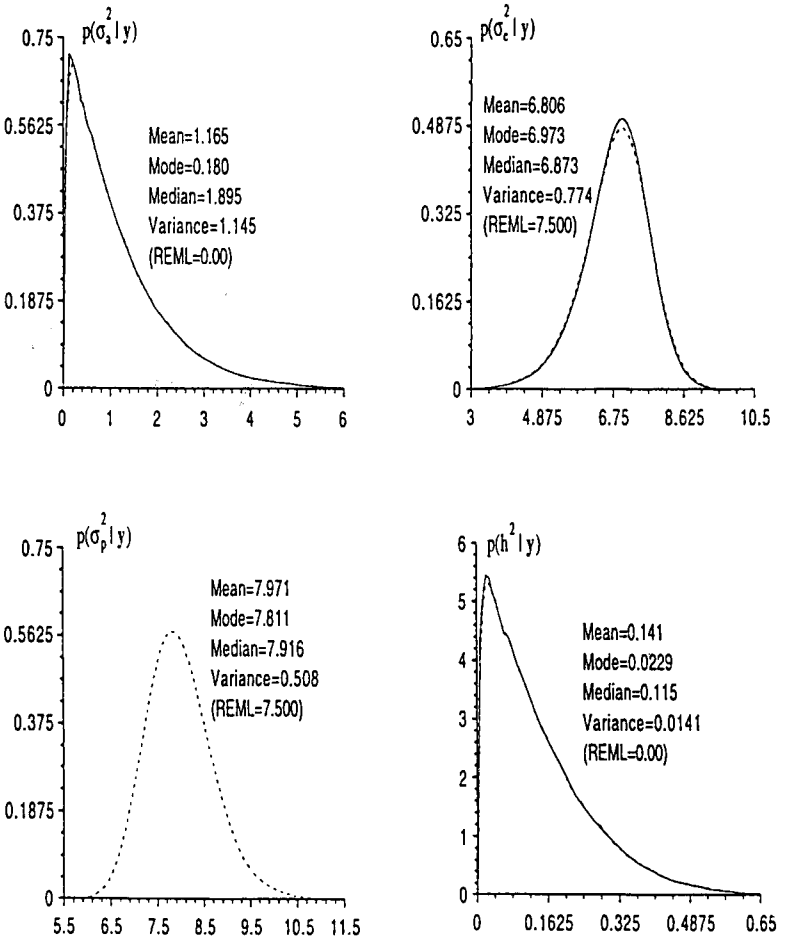

Fig 1. Estimated marginal posterior densities of variance components (σ_a^2, σ_e^2), total variance (σ_p^2) and heritability (h^2) using PART data (generations 0–2) simulated with $h^2 = 0.2$. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2).

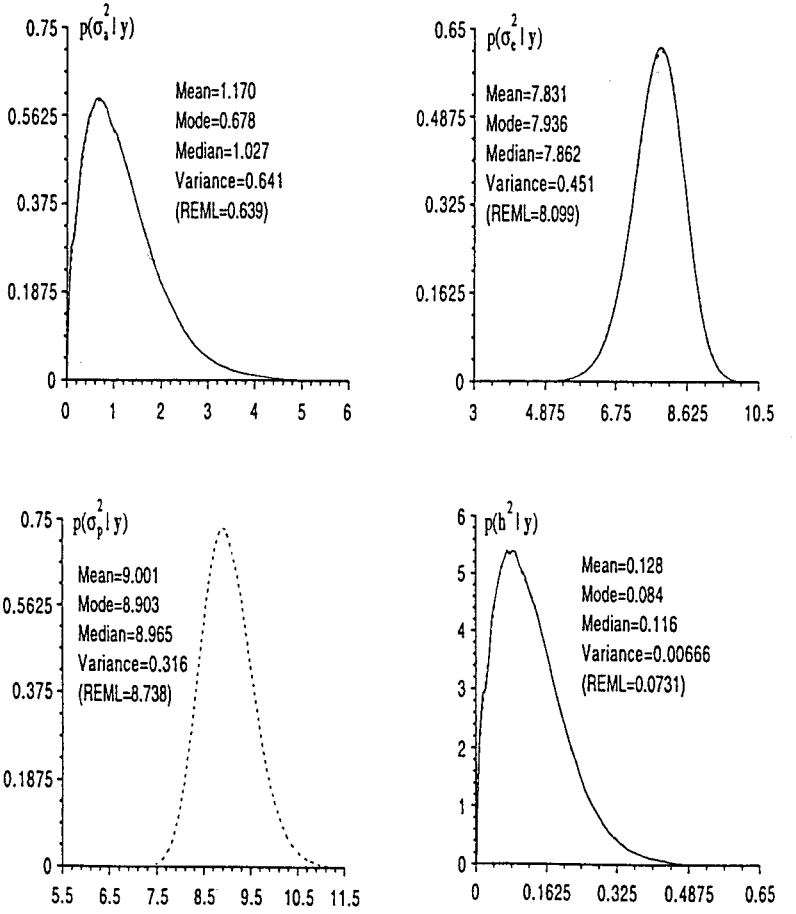

Fig 2. Estimated marginal posterior densities of variance components (σ_a^2, σ_e^2), total variance (σ_p^2) and heritability (h^2) using WHOLE data (generations 0–5) simulated with $h^2 = 0.2$. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2).

For the data simulated with $h^2 = 0.2$, the posterior distributions reflected considerable uncertainty about the values of the parameters. As expected, the variances of the posterior distributions were smaller in WHOLE (fig 2) than in PART (fig 1). In particular, the PART analysis of σ_a^2 and h^2 showed highly skewed distributions; the mean, the mode and the median differed from each other. The REML estimates in the PART analysis for the additive genetic variance and heritability were very close to zero. Asymptotic confidence intervals were not computed, but they would probably include a region outside the parameter space. The Bayesian analysis indicated with considerable probability that both parameters are non-zero. When WHOLE was considered (fig 2), the degree of uncertainty

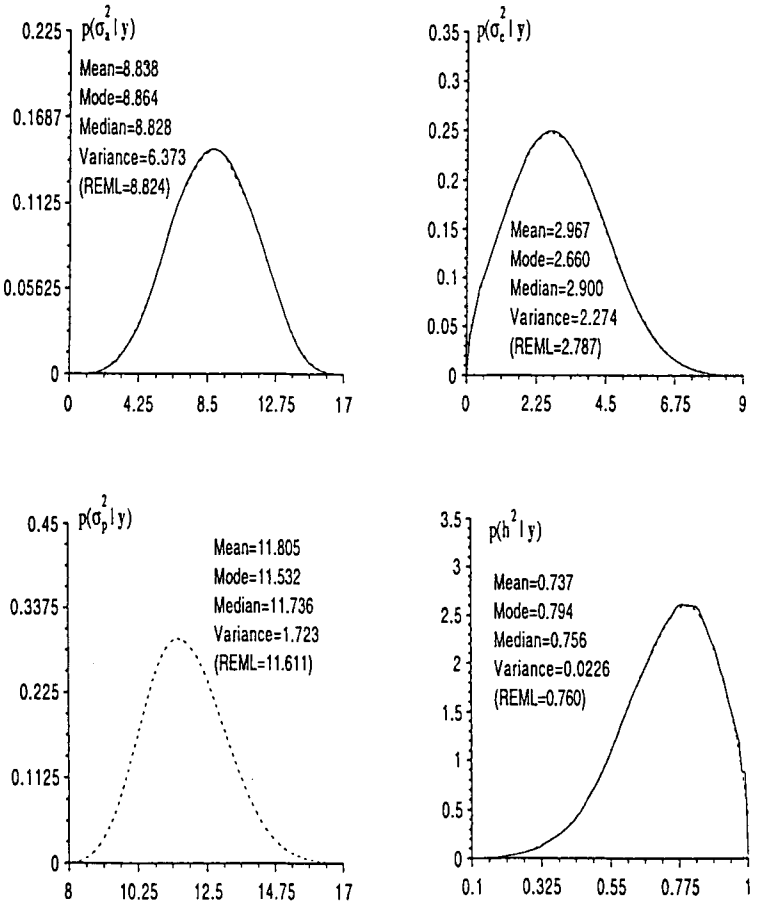

Fig 3. Estimated marginal posterior densities of variance components (σ_a^2, σ_e^2), total variance (σ_p^2) and heritability (h^2) using PART data (generations 0–2) simulated with $h^2 = 0.5$. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2).

was reduced, but point estimates in the Bayesian analysis of genetic variance and of heritability were different from those in the REML estimates. For example, the posterior mean, mode and median of heritability were 0.13, 0.08 and 0.12, respectively, in comparison with the REML estimate of 0.07. The lack of symmetry of the posterior distribution of σ_a^2 and h^2 clearly suggests that the simulated selection experiment does not have a high degree of resolution concerning inferences about genetic parameters. This point would not be as forcefully illustrated in a standard REML analysis, where, at best, one would get joint approximate asymptotic confidence intervals for the parameters of interest.

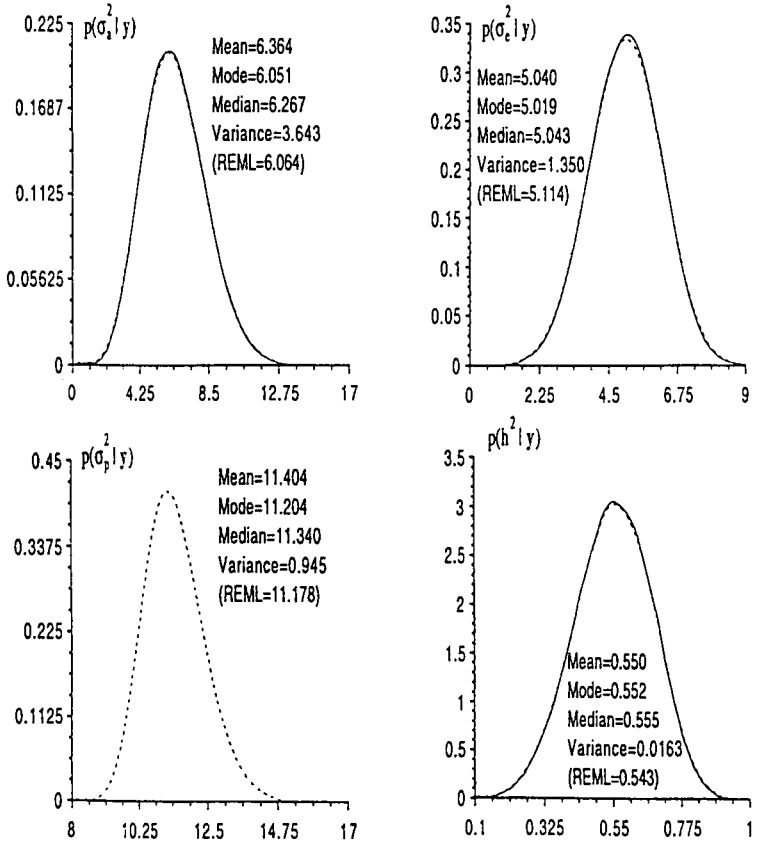

Fig 4. Estimated marginal posterior densities of variance components (σ_a^2, σ_e^2), total variance (σ_p^2) and heritability (h^2) using WHOLE data (generations 0–5) simulated with $h^2 = 0.5$. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2).

In the case of $h^2 = 0.5$ (fig 3 and 4), it should be noted that the distributions were less skewed than for $h^2 = 0.2$, although still skewed for PART (fig 3). Both the Bayesian and the REML analyses suggested a moderate to high heritability, but the fact that the posterior distribution of heritability for PART (fig 3) was very skewed indicates that more data should be collected for accurate inferences about heritability. The WHOLE analysis (fig 4) yielded nearly symmetric posterior distributions of heritability with smaller variance, centered at 0.55 (the REML estimate was 0.543). It should be noted that the residual variance was poorly estimated in PART (fig 3), but the Bayesian analysis clearly depicted the extent of uncertainty about this parameter. This also illustrates the fact that a variance component smaller in value relative to others in the model is harder to estimate,

contrary to the common belief in animal breeding that residual variance is always well estimated.

The assessment of response to selection for the population simulated with $h^2 = 0.2$ is given in figures 5 and 6 for the PART and WHOLE data analyses, respectively. The true response in the simulated data at a given generation, computed as the average of true genetic values within the generaton in question, was $TR = 0.643$ units and $TR = 1.783$ units, for PART and WHOLE. The regressions of true response on generation were 0.321 and 0.343, for PART and WHOLE, respectively. The hypothesis of no response to selection cannot be rejected in the PART analysis (fig 5). The classical approach gave an estimated response of zero. However, the Bayesian analysis documents the extent of uncertainty, and assigns considerable posterior probability to the proposition that selection may be effective; the posterior probability of response (*eg*, $\sigma_1 > 0$ or $TR = t_2 - t_0 > 0$) was much larger than that of the complement. The strength of additional evidence brought up in the whole analysis (fig 6) supports the hypothesis that selection was effective. For example, the difference in genetic mean between generation 0 and 5 was centered around 1.852, though the variance of the posterior distribution was as large as 1.075. The classical analysis gave an estimated value of $ST = 1.296$ units, but no exact measure of variation can be associated to this estimate. Approximate results could be obtained, for example, using Monte-Carlo simulations, but this was not attempted in the present work.

In the simulation with $h^2 = 0.5$, the true response in the simulated data, computed as the average of true genetic values from the relevant generations, was $TR = 3.258$ units and $TR = 7.148$ units for the PART and WHOLE data sets, respectively. The associated regressions on generation were 1.629 and 1.436, respectively. The Bayesian and classical analyses indicated response to selection beyond reasonable doubt (fig 7 and 8). Note, however, that the posterior distribution of $TR = t_2 - t_0$ in PART was skewed (fig 7). In figure 8, all the posterior distributions were nearly symmetric, and the classical and Bayesian analyses were essentially in agreement. In spite of this symmetry, considerable uncertainty remained about the true rate of response to selection. For example, values of $t_5 - t_0$ ranging from 3 to 11 units have appreciable posterior density (fig 8). In the light of this, results from a similar experiment where realized genetic change is, say, 50% higher or lower than our posterior mean of about 6.7 units, could not be interpreted as yielding conflicting evidence.

In order to study the influence of different priors on inferences about heritability and response to selection, part of the above analyses using the same data, was repeated assuming that variance components followed, *a priori*, the inverse chi-square distributions [9]. The 'degree of belief' parameters ν_i and the *a priori* value of the variance components S_i^2, were assessed by the method of moments from an independently simulated data set with base population heritability of 0.5, and consisting of 3 cycles of selection. An REML estimate of heritability from this data was 0.46 with an asymptotic variance of 2.94×10^{-2}. The method of moments as described in Wang *et al* (1994b), produced the following estimates: $\nu_a = 11$; $\nu_e = 464$; $S_a^2 = 4.50$; $S_e^2 = 5.19$. A reanalysis of the PART and WHOLE replicate (heritability = 0.5) yielded the results shown in table I. The figures in the table clearly illustrate how increasing the amount of information in the data modifies

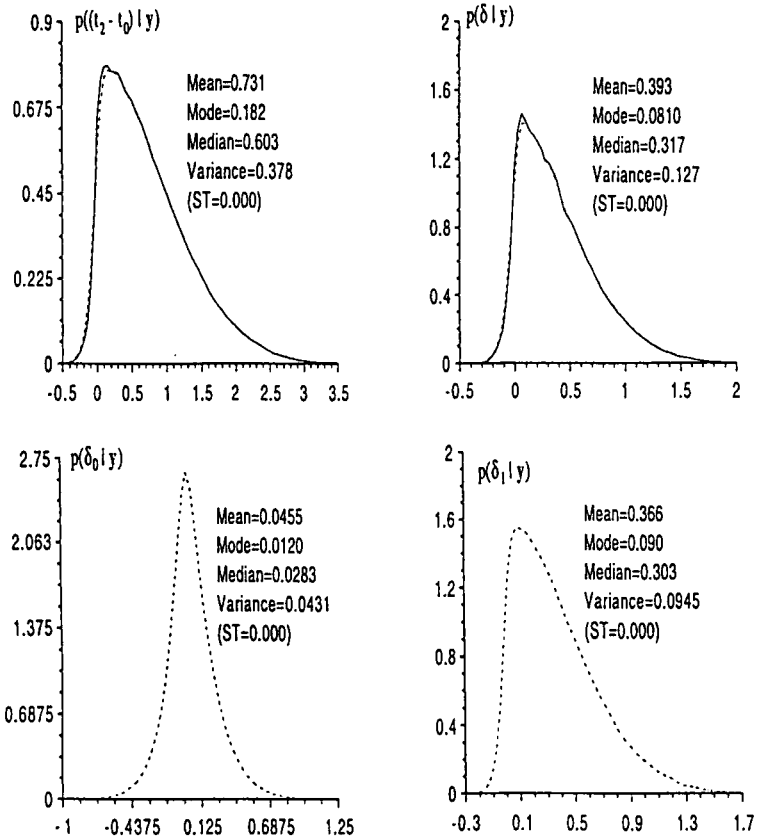

Fig 5. Estimated marginal posterior densities of response to selection using PART data (generations 0–2) simulated with heritability = 0.2: total response ($TR = t_2 - t_0$); regression of additive genetic values on generation through origin (δ); and intercept (δ_0) and slope (δ_1) of linear regression of additive genetic values on generation. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2). *ST*: response from classical (BLUP) analysis.

Table I. Expected value E(·) and variance (Var) of the marginal posterior distributions of heritability (h^2) and of total response to selection (TR) for the PART and WHOLE data, using as priors either uniform distributions (UNIFORM) or inverted χ^2 distributions (INVCHISQ).

	PART		*WHOLE*	
	UNIFORM	*INVCHISQ*	*UNIFORM*	*INVCHISQ*
E(h^2)	0.737	0.501	0.550	0.529
Var × 10^2	2.26	0.293	1.63	0.235
E(TR)	3.49	2.44	6.698	6.534
Var	0.733	0.332	2.836	0.907

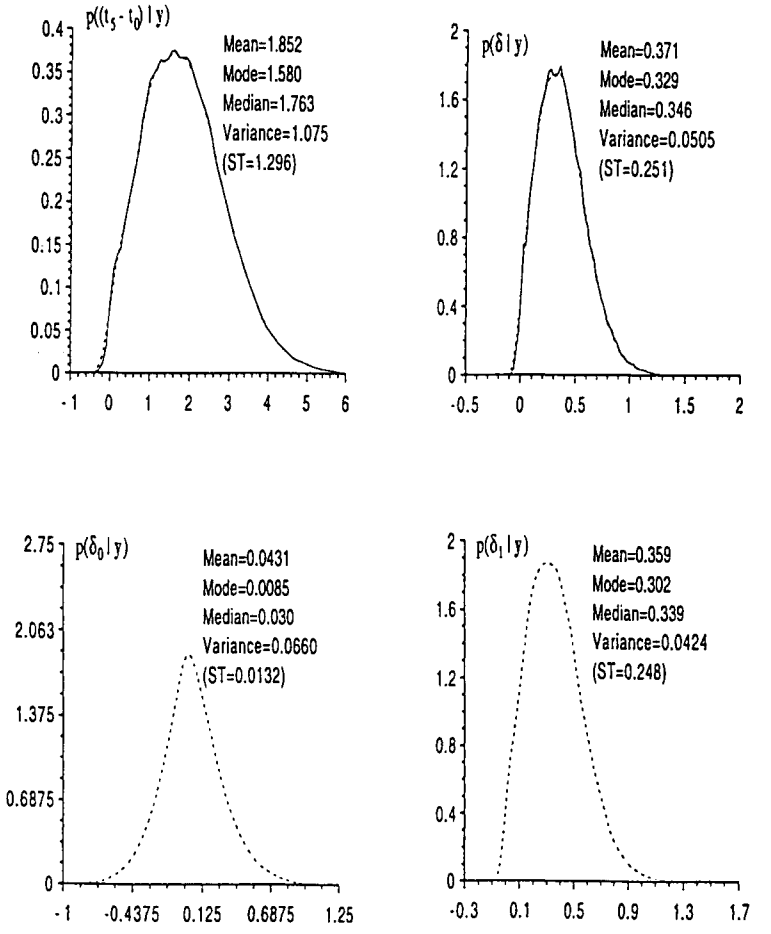

Fig 6. Estimated marginal posterior densities of response to selection using WHOLE data (generations 0–5) simulated with heritability = 0.2: total response ($TR = t_5 - t_0$); regression of additive genetic values on generation through origin (δ); and intercept (δ_0) and slope (δ_1) of linear regression of additive genetic values on generation. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2). *ST*: response from classical (BLUP) analysis.

the input contributed by the prior. Thus in the PART dataset, the uniform prior and the informative prior (which implies a prior value of heritability of 0.46) lead to clear differences in the expected value of the marginal posterior distribution of heritability and selection response. However, in the WHOLE data set, inferences using either prior are very similar. The figures also illustrate how the variance of marginal posterior distributions is reduced when an informative prior is used.

The examples indicate the power of the Bayesian analysis to reveal uncertainty in response to selection when the information contained in the data about the

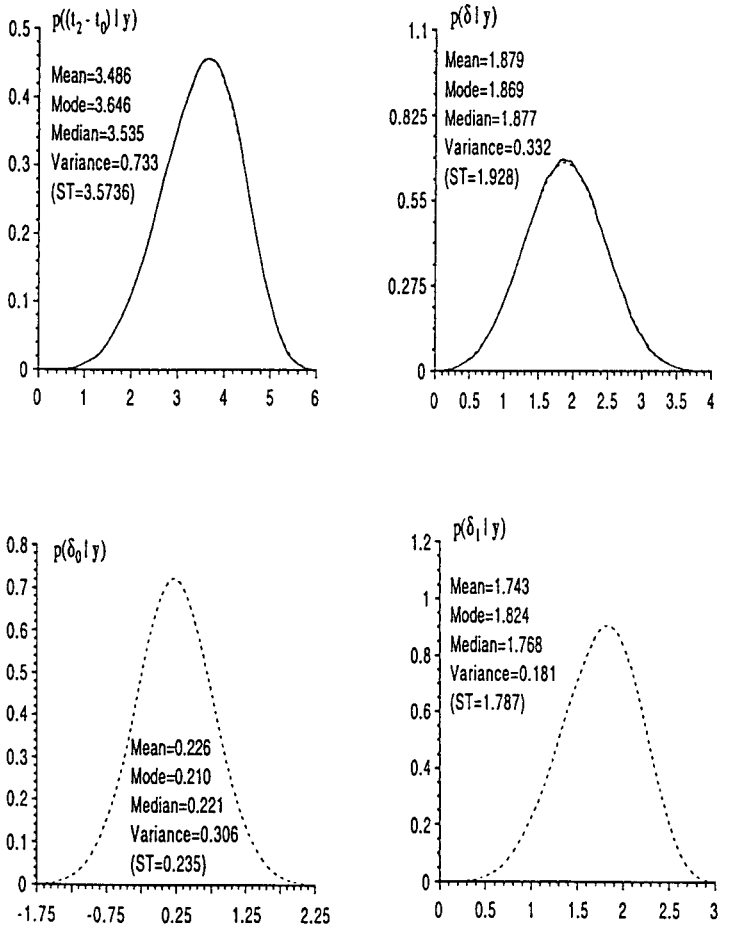

Fig 7. Estimated marginal posterior densities of response to selection using PART data (generations 0–2) simulated with heritability = 0.5: total response ($TR = t_2 - t_0$); regression of additive genetic values on generation through origin (δ); and intercept (δ_0) and slope (δ_1) of linear regression of additive genetic values on generation. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2). *ST*: response from classical (BLUP) analysis.

appropriate parameters is small (*eg*, PART analysis at $h^2 = 0.2$). Other plots, such as marginal posterior distribution of generation means, for example, are possible and could illustrate the evolution of the drift variance as the experiment progresses.

DISCUSSION

We have presented a way of analysing response to selection in experimental or non-experimental settings, from a Bayesian viewpoint. The end-point of the analysis

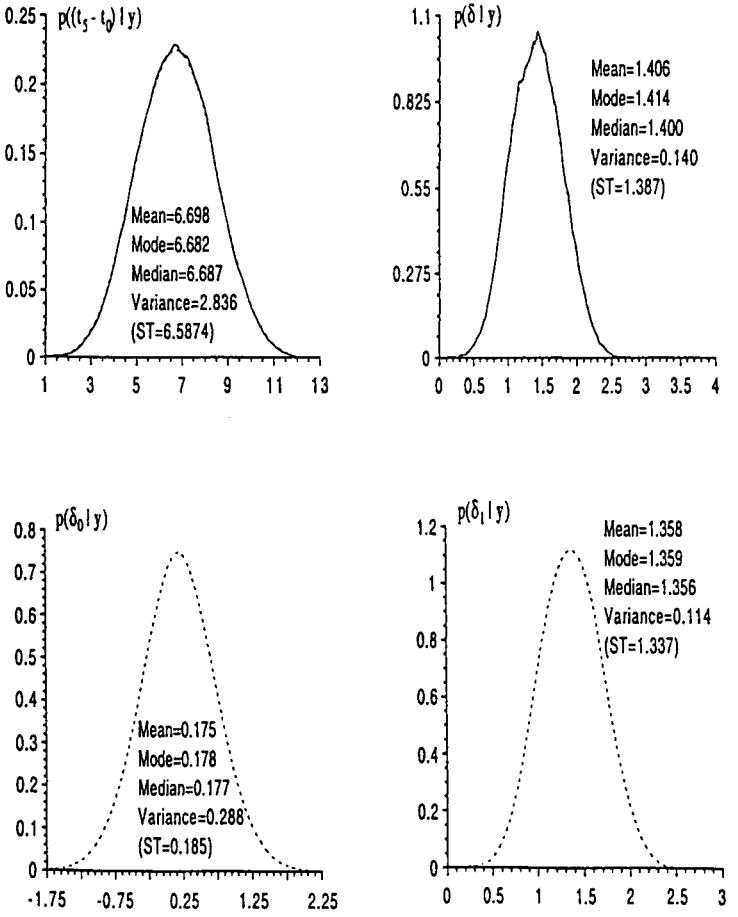

Fig 8. Estimated marginal posterior densities of response to selection using WHOLE data (generations 0–5) simulated with heritability = 0.5: total response ($TR = t_5 - t_0$); regression of additive genetic values on generation through origin (δ); and intercept (δ_0) and slope (δ_1) of linear regression of additive genetic values on generation. Solid curves are based on conditional density estimator (3); dotted curves are based on the normal kernel density estimator (2). *ST*: response from classical (BLUP) analysis.

is the construction of a marginal posterior distribution of a measure of selection response which, in this study, was defined as a linear function of additive genetic values. The marginal distribution can be viewed as a weighted average of an infinite number of conditional distributions, where the weighting function is the marginal posterior distribution of variances and other parameters, which are not necessarily of interest. The approach relies heavily on the result (Gianola and Fernando, 1986) that if the data on which selection was based are included in the analysis, the Bayesian approach accounts for selection automatically, in the sense

that the joint posterior or any marginal posterior distribution is the same with or without selection. This is a particular case of what has been known as 'ignorable selection' (Little and Rubin, 1987; Im *et al*, 1989) and it implies that selection can be ignored if the information of any data-based selection is contained in the Bayesian model. Furthermore, inferences pertain to parameters, *eg*, heritability, of the base population.

As shown with the simulated data, the marginal posterior distributions of variances and functions thereof are often not symmetric. This fact is taken into account when computing the marginal posterior distributions of measures of selection response. In this way, the uncertainty associated with all nuisance parameters in the model is taken into account when drawing inferences about response to selection. This is in marked contrast with the estimation of response that is obtained using BLUP with an animal model, which assumes the variance ratio known, and gives 100% weight to an estimate of this ratio.

The analysis of the simulated data used uniform distributions for the fixed effects and variance components. The choice of appropriate prior distributions is a contentious issue in Bayesian inference, especially when these priors are supposed to convey vague initial knowledge. Injudicious choice of noninformative prior distributions can lead to improper posterior distributions, and this may not be easy to recognize, especially when the analysis is based on numerical methods, as in the present work. The subject is still a matter of debate and an important concept is that of reference priors (Bernardo, 1979; Berger and Bernardo, 1992), which have the property that they contribute with minimum information to the posterior distribution while the information arising from the likelihood is maximized. In the single parameter case, reference priors often yield the Jeffreys priors (Jeffreys, 1961). The uniform prior distributions we have chosen for the fixed effects and the variance components represent a state of little prior knowlege, but are clearly not invariant under transformations and must be viewed as simple *ad hoc* approximations to the appropriate reference prior distributions. We have illustrated, however, that when the amount of information contained in the data is adequate inferences are affected little by the choice of priors. It is not generally a simple matter to decide when one is in a setting of adequate information, and it may be therefore revealing to carry out analyses with different priors to study how inferences are affected (Berger, 1985). If use of different priors leads to very different results, this indicates that the information in the likelihood is weak and more data ought to be collected in order to draw firmer conclusions.

The richness of the information that can be extracted using the Bayesian approach is at the cost of computational demands. The demand is in terms of computer time rather than in terms of programming complexity. However, an important limitation of iterative simulation methods is that drawing from the appropriate posterior density takes place only in the limit, as the number of draws becomes infinite. It is not easy to check for convergence to the correct distribution, and there is little developed theory indicating how long the Gibbs chain should be. The rate of convergence can be exceedingly slow, especially when random variables are highly correlated, which is the case with animal models. Further, there is often a possibility that the Gibbs sequence may get 'trapped' near zero which is an absorbing state, in which case one must reinitialize the chain (Zeger and Karim,

1991). Gibbs sampling convergence is an area of active research where a number of partial answers based on pragmatic approaches have been suggested. These include checking for serial correlations, monitoring the behavior of several series of runs of the chains starting from a wide range of initial values and checking both within- and between-series variation (Gelman and Rubin, 1992). Another possibility, given a Gibbs sample of size m, is to vary the length of the sequence and to overlay plots of the estimated densities to see if these are distinguishable (Gelfand *et al*, 1990). Convergence can probably be accelerated if correlated random variables (such as additive genetic values in animal models) are blocked together, at the expense of having to sample from multivariate conditional distributions (Smith and Roberts, 1993). In general though, selection experiments are scarce and expensive, and the cost of the additional computation needed for carrying out the Bayesian analysis in often marginal relative to the whole cost.

The great appeal of this method is that it yields a Monte-Carlo estimate of a full marginal posterior distribution of a parameter of interest, from which probabilities that the parameter lies between specified values can be easily computed. This is particularly relevant in the case where the asymptotic normality of posterior distributions is difficult to justify, which can often be the case in selection experiments. The pictorial representation of marginal posterior distributions, and the associated possibility of making precise probability statements, provide for a very rich inferential framework. Errors incurred when estimating a posterior density by Monte-Carlo methods can be made arbitrarily small by increasing the chain length in the Gibbs sampling, at least in principle.

The Bayesian analysis can also be useful to study the design of selection experiments. The literature on experimental design relies heavily on assumptions such as absence of nuisance parameters and knowledge of base population parameters. Studies about the use of animals models in selection experiments have concentrated on analysis rather than design issues. This is partly because in a frequentist setting it is difficult to compute the sampling variance of the estimator of response, especially when variances used as priors have been estimated from the data at hand. Using the Bayesian approach, a variety of designs could be studied and their efficiency compared by means of analyses of predictive distributions.

ACKNOWLEDGMENTS

The topic of this research emerged through conversations which one of us (DAS) held with V Ducrocq (INRA). WG Hill (Edinburgh) suggested the study of the influence of different priors on inferences about response to selection, and JM Bernardo (Valencia) suggested parameterizing the priors for fixed effects and variance components as proper uniform distributions. Both made insightful comments on an earlier draft of this paper. Part of this research was financed by the Danish Agricultural and Veterinary Research Council.

REFERENCES

Blair HT, Pollak EJ (1984) Estimation of genetic trend in a selected population with and without the use of a control population. *J Anim Sci* 58, 878-886

Berger JO (1985) *Statistical Decision Theory and Bayesian analysis.* Springer-Verlag

Berger JO, Bernardo JM (1992) Reference priors in a variance components problem. *In: Proc of the Indo-USA Workshop on Bayesian Analysis in Statistics and Econometrics.* Springer-Verlag

Bernardo JM (1979) Reference posterior distributions for Bayesian inference (with discussion). *J R Stat Soc Series B* 41, 113-147

Casella G, George EI (1992) Explaining the Gibbs sampler. *Am Stat* 46, 167-174

Fernando RL, Gianola D (1990) Statistical inferences in populations undergoing selection or nonrandom mating. *In: Advances in Statistical Methods for Genetic Improvement of Livestock* (D Gianola, K Hammond, eds) Springer-Verlag, 437-453

Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 85, 398-409

Gelfand A, Hills SE, Racine-Poon A, Smith AFM (1990) Illustration of Bayesian inference in normal data models using Gibbs sampling. *J Am Stat Assoc* 85, 972-985

Gelfand A, Smith AFM, Lee TM (1992) Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *J Am Stat Assoc* 87, 523-532

Gelman A, Rubin DB (1992) A single series from the Gibbs sampler provides a false sense of security. *In: Bayesian Statistics IV* (JM Bernardo, JO Berger, AP Dawid, AFM Smith, eds) Oxford University Press, 625-631

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721-741

Geyer CJ (1992) Practical Markov chain Monte Carlo (with discussion). *Stat Sci* 7, 467-511

Gianola D, Fernando RL (1986) Bayesian methods in animal breeding. *J Anim Sci* 63, 217-244

Gianola D, Foulley JL, Fernando RL (1986) Prediction of breeding values when variances are not known. *Genet Sel Evol* 18, 485-498

Gianola D, Fernando RL, Im S, Foulley JL (1989) Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. *Genome* 31, 768-777

Henderson CR (1973) Sire evaluation and genetic trends. *In: Proc Anim Breed and Genet Symp in honor of Dr JL Lush.* ASAS and ADSA, Champaign

Henderson CR (1976) A simple method for computing the inverse of a numerator relation matrix used in prediction of breeding values. *Biometrics* 32, 69-83

Im S, Fernando RL, Gianola D (1989) Likelihood inferences in animal breeding under selection: a missing data theory view point (with discussion). *Genet Sel Evol* 21, 399-414

IMSL, Inc (1989) *IMSL STAT/LIBRARY: FORTRAN* Subroutines for Statistical Analysis

Jensen J, Wang CS, Sorensen DA, Gianola D (1994) Marginal inferences of variance and covariance components for traits influenced by maternal and direct genetic effects using the Gibbs sampler. *Acta Agric Scand* (in press)

Jeffreys H (1961) *Theory of Probability.* Clarendon Press, Oxford (3rd edition)

Kackar RN, Harville DA (1981) Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. *Commun Stat Theor Math* A10, 1249-1261

Lindley DV, Smith AFM (1972) Bayes estimates for the linear model. *J R Stat Soc Series B*, 34, 1-18

Little RJA, Rubin DB (1987) *Statistical Analysis with Missing Data.* John Wiley & Sons, Inc, New York

Meyer K, Hill WG (1991) Mixed-model analysis of a selection experiment for food intake in mice. *Gen Rest* 57, 71-81

Quaas RL (1976) Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* 32, 949-953

Silverman BW (1986) *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London

Smith AFM, Roberts GO (1993) Bayesian computation *via* the Gibbs sampler and related Markov chain Monte-Carlo methods (with discussion). *J R Stat Soc Series B* 55, 3-23

Sorensen DA, Kennedy BW (1986) Analysis of selection experiments using mixed-model methodology. *J Anim Sci* 63, 245-258

Sorensen DA, Johansson K (1992) Estimation of direct and correlated responses to selection using univariate animal models. *J Anim Sci* 70, 2038-2044

Thompson R (1986) Estimation of realized heritability in a selected population using mixed-model methods. *Genet Sel Evol* 18, 475-483

Wang CS, Rutledge JJ, Gianola D (1993) Marginal inferences about variance components in a mixed linear model using Gibbs sampling. *Genet Sel Evol* 21, 41-62

Wang CS, Rutledge JJ, Gianola D (1994a) Bayesian analysis of mixed linear models *via* Gibbs sampling with an application to litter size in Iberian pigs. *Genet Sel Evol* 26, 91-115

Wang CS, Gianola D, Sorensen DA, Jensen J, Cristensen A, Rutledge JJ (1994b) Response to selection for litter size in Danish Landrace pigs: A Bayesian analysis. *Theor Appl Genet*, in press

Zeger SL, Karim MR (1991) Generalized linear models with random effects: a Gibbs sampling approach. *J Am Stat Assoc* 86, 79-86