

Analysis of a genetically structured variance heterogeneity model using the Box–Cox transformation

YE YANG, OLE F. CHRISTENSEN AND DANIEL SORENSEN*

Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Aarhus University, DK-8830 Tjele, Denmark

(Received 5 July 2010 and in revised form 8 September 2010; first published online 25 February 2011)

Summary

Over recent years, statistical support for the presence of genetic factors operating at the level of the environmental variance has come from fitting a genetically structured heterogeneous variance model to field or experimental data in various species. Misleading results may arise due to skewness of the marginal distribution of the data. To investigate how the scale of measurement affects inferences, the genetically structured heterogeneous variance model is extended to accommodate the family of Box–Cox transformations. Litter size data in rabbits and pigs that had previously been analysed in the untransformed scale were reanalysed in a scale equal to the mode of the marginal posterior distribution of the Box–Cox parameter. In the rabbit data, the statistical evidence for a genetic component at the level of the environmental variance is considerably weaker than that resulting from an analysis in the original metric. In the pig data, the statistical evidence is stronger, but the coefficient of correlation between additive genetic effects affecting mean and variance changes sign, compared to the results in the untransformed scale. The study confirms that inferences on variances can be strongly affected by the presence of asymmetry in the distribution of data. We recommend that to avoid one important source of spurious inferences, future work seeking support for a genetic component acting on environmental variation using a parametric approach based on normality assumptions confirms that these are met.

1. Introduction

The classical model of quantitative genetics assumes that genotypes differ in mean but the environmental variance (conditional variance of phenotype, given genotype) is homogeneous across genotypes (this model is referred to as HOM hereinafter). The assumption of homogeneity may not be valid and models with non-genetically structured departures from variance homogeneity were introduced in the 1990s (Foulley *et al.*, 1992; Gianola *et al.*, 1992; San Cristobal *et al.*, 1993). In particular, Foulley & Quaas (1995) proposed models of heterogeneity for both residual and other components of variance.

A significant extension of the classical genetic model, with implications in breeding and evolutionary studies, posits that both mean and variability differ between genotypes (San Cristobal-Gaudy *et al.*, 1998).

(The extended model is labelled HET hereinafter). In animal and plant breeding, the focus is to change the mean of a trait by selection; the HET model opens the possibility to reduce variation by selection leading to more homogeneous products (e.g. Hill & Zhang, 2004; Mulder *et al.*, 2007, 2008). In evolutionary biology, most of the models developed to study the maintenance of phenotypic variation assume that environmental variance is constant and explain the levels of variation by invoking a balance between the gain of genetic variance by mutation and its loss by different forms of selection and drift. Zhang & Hill (2005) discuss models where environmental variance is partly under genetic control and study conditions for its maintenance under stabilizing selection.

Early evidence for a genetic component affecting environmental variation stems from comparisons of levels of variation between inbred lines and the F_1 cross between them, with inbreds showing in general larger variance (Falconer & Mackay, 1996). More

* Corresponding author. email: Sorensen@humo.dk

recently, experimental evidence from isogenic chromosome substitution lines in *Drosophila* was presented by Mackay & Lyman (2005). Using genetic marker information, support for genetic variance heterogeneity has been reported also by Ordas *et al.*, (2008) in maize and by Ansel *et al.* (2008) in yeast. Ansel *et al.* (2008) work with isogenic yeast cells and provide convincing evidence for heterogeneity of gene expression between genotypes, and identify three quantitative trait loci (QTLs) involved in its control.

With the exception of the limited experimental evidence specifically designed to test for genetic control of residual variability, support for the presence of genes acting on environmental variation can be sought fitting the HET model to data and studying the quality of its fit using modern computation tools. Evidence of this kind has been reported using litter size in pigs (Sorensen & Waagepetersen, 2003), adult weight in snails (Ros *et al.*, 2004), uterine capacity in rabbits (Ibáñez *et al.*, 2008), body weight in poultry (Rowe *et al.*, 2006; Wolc *et al.*, 2009), slaughter weight in pigs (Ibáñez *et al.*, 2007) and litter size and weight at birth in mice (Gutierrez *et al.*, 2006).

In all the cases where the quality of fit of the HET and HOM models was compared, various statistical criteria favoured the HET model. An important issue is to clarify whether these results are an artefact of the scale of measurement. Under the HET model, the skewness of the marginal distribution of the data is directly proportional to the correlation between additive genetic values affecting mean and variance (Ros *et al.*, 2004). However skewness can arise in other ways and if it does it can cause the HET model to fit better than the HOM model, leading to the spurious conclusion that genetic heterogeneity is present, when it is absent. Alternatively, genetic heterogeneity of variance may be present, in which case skewness from other causes may affect inferences about the correlation between genetic effects on mean and variance.

These problems can be partly addressed finding the appropriate scale on which to analyse the data using the Box–Cox transformation parameter (Box & Cox, 1964). A large body of literature dealing with transformations has developed following the classical paper of Box & Cox (1964). Early work using transformations in animal breeding are Ibe & Hill (1988) and Gianola *et al.* (1990). Compelling reasons for using the transformed model have typically been the need to achieve additivity of parameters at the level of the mean, constancy and homogeneity of the error variance and normality of the distribution of observations, although a transformation may not achieve all these goals simultaneously. Here, a transformation is sought that induces normality and linearity at the level of the conditional distribution of data. This is an

important modelling assumption that must be fulfilled to avoid distorted inferences.

The main objective of this work is to investigate whether an analysis performed under the ‘appropriate’ scale still supports the HET model, or whether simpler models are to be preferred. The Box–Cox transformation parameter is regarded as an unknown to be inferred from the data, jointly, with the remaining parameters of the model. The model is fitted to data on litter size in pigs and rabbits. In order to obtain insight into the quality of inferences about the transformation parameter when it is fitted simultaneously with the remaining parameters, including the coefficient of correlation, a simulation study involving sets of data that differ in the number of repeated records per individual is also conducted.

Implementation of the model is via McMC (Markov chain Monte Carlo) and two steps are involved. The first step aims at finding the modal value of the marginal posterior distribution of the transformation parameter. In the second step, the model is fitted conditional on this modal value. Several McMC strategies were investigated and the reported results are based on a fixed scan hybrid algorithm that resulted in the best mixing behaviour.

The article is organized as follows: section 2 introduces the Box–Cox model with genetically structured variance heterogeneity and discusses the prior specifications and McMC implementation. Section 3 describes briefly the simulated, the rabbit and the pig litter size data. Section 4 contains the outcome of applying the methodology to the simulated data and to the rabbit and pig litter size data. A discussion is provided in section 5.

2. Methods

(i) Box–Cox model with genetically structured variance heterogeneity

The proposed model, labelled BCHET hereinafter, combines the HET model with a Box–Cox transformation parameter. Let $\mathbf{y} = (y_i)_{i=1}^n$ be the vector of records, which are assumed to be strictly positive. The model states that conditionally on the vectors of location and dispersion parameters, for some unknown transformation parameter λ , the distribution of the transformed vector of phenotypic data $\mathbf{y}^{(\lambda)} = (y_i^{(\lambda)})_{i=1}^n$ satisfies the normal theory assumptions, that is,

$$\mathbf{y}^{(\lambda)} | \lambda, \mu, \mu^*, \mathbf{b}, \mathbf{a}, \mathbf{p}, \mathbf{b}^*, \mathbf{a}^*, \mathbf{p}^* \sim N(\mathbf{1}\mu + \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{p}, \mathbf{R}). \quad (1)$$

Here \mathbf{R} , the conditional variance of the sampling model, is the diagonal matrix with $\exp\{1\mu^* + \tilde{\mathbf{X}}\mathbf{b}^* + \mathbf{Z}\mathbf{a}^* + \mathbf{W}\mathbf{p}^*\}$ on its diagonal, and is interpreted as the environmental variance. The transformation is

defined as follows (Box & Cox, 1964):

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & (\lambda \neq 0), \\ \log y & (\lambda = 0). \end{cases} \quad (2)$$

Note that here the transformation parameter λ is introduced to guarantee normality of the conditional distribution and linearity of the conditional expectation, but not to achieve constant variance. Indeed, \mathbf{R} in (1) is heterogeneous due to the presence of $\tilde{\mathbf{X}}\mathbf{b}^*$, \mathbf{a}^* and \mathbf{p}^* ; the last two terms being specific to each individual.

Matrices \mathbf{X} , $\tilde{\mathbf{X}}$, \mathbf{Z} and \mathbf{W} are known incidence matrices, the scalar μ is the mean, the vector \mathbf{b} contains the regression coefficients, the scalar $\exp(\mu^*)$ is the mean environmental variance, while the vector $\exp(\mathbf{b}^*)$ contains the environment-specific scaling effects. The vectors of genetic effects $\mathbf{a} = (a_i)_{i=1}^{n_1}$ and $\mathbf{a}^* = (a_i^*)_{i=1}^{n_1}$ are assumed to jointly follow a multivariate normal distribution

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{a}^* \end{pmatrix} | \mathbf{G} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{G} \otimes \mathbf{A} \right),$$

where \mathbf{A} is the known additive relationship matrix and \mathbf{G} is a 2×2 covariance matrix

$$\mathbf{G} = \begin{pmatrix} \sigma_a^2 & \rho\sigma_a\sigma_{a^*} \\ \rho\sigma_a\sigma_{a^*} & \sigma_{a^*}^2 \end{pmatrix}. \quad (3)$$

In equation (3), ρ is the coefficient of genetic correlation between \mathbf{a} and \mathbf{a}^* , and σ_a^2 and $\sigma_{a^*}^2$ are the additive genetic variances at the level of the mean and environmental variance, respectively. Vectors $\mathbf{p} = (p_i)_{i=1}^{n_2}$ and $\mathbf{p}^* = (p_i^*)_{i=1}^{n_2}$ represent permanent environmental effects (on mean and on variance) and are assumed to be independently normally distributed with

$$\mathbf{p} | \sigma_p^2 \sim N(0, \sigma_p^2 \mathbf{I}),$$

$$\mathbf{p}^* | \sigma_{p^*}^2 \sim N(0, \sigma_{p^*}^2 \mathbf{I}),$$

where \mathbf{I} is the identity matrix, and σ_p^2 and $\sigma_{p^*}^2$ are the permanent environmental variances at the level of the mean and environmental variance, respectively.

The coefficient of skewness of the marginal distribution of the datum $y_i^{(\lambda)}$ (marginalized over \mathbf{a} , \mathbf{a}^* , \mathbf{p} , \mathbf{p}^*) under BCHET is (Ros *et al.*, 2004)

$$\begin{aligned} & \frac{E[(y_i^{(\lambda)} - E(y_i^{(\lambda)}))^3] | \lambda, \mu, \mu^*, \mathbf{b}, \mathbf{b}^*, \sigma_a^2, \sigma_{a^*}^2, \rho, \sigma_p^2, \sigma_{p^*}^2]}{\text{Var}(y_i^{(\lambda)} | \lambda, \mu, \mu^*, \mathbf{b}, \mathbf{b}^*, \sigma_a^2, \sigma_{a^*}^2, \rho, \sigma_p^2, \sigma_{p^*}^2)^{3/2}} \\ &= \rho \frac{3\sigma_a\sigma_{a^*}\exp(\mu^* + (\tilde{\mathbf{X}}\mathbf{b}^*)_i + \frac{1}{2}\sigma_a^2 + \frac{1}{2}\sigma_{p^*}^2)}{\left(\exp(\mu^* + (\tilde{\mathbf{X}}\mathbf{b}^*)_i + \frac{1}{2}\sigma_a^2 + \frac{1}{2}\sigma_{p^*}^2) + \sigma_a^2 + \sigma_{p^*}^2\right)^{3/2}}, \end{aligned}$$

indicating that the sign of the coefficient of skewness is governed by ρ .

(ii) Conditional likelihood function

The conditional likelihood is proportional to the probability density function of the data given the parameters, including \mathbf{a} , \mathbf{a}^* , \mathbf{p} , \mathbf{p}^* . Written in terms of the original untransformed observations, from (1), this leads to

$$|\mathbf{R}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y}^{(\lambda)} - E(\mathbf{y}^{(\lambda)}))' \mathbf{R}^{-1} (\mathbf{y}^{(\lambda)} - E(\mathbf{y}^{(\lambda)})) \right\} J(\lambda; \mathbf{y}), \quad (4)$$

where

$$E(\mathbf{y}^{(\lambda)}) = 1\mu + \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{W}\mathbf{p}, \quad (5)$$

\mathbf{R} and $\mathbf{y}^{(\lambda)}$ have been defined in connection with (1), and the Jacobian of the transformation is

$$\begin{aligned} J(\lambda; \mathbf{y}) &= \prod_{i=1}^n \left| \frac{\partial y_i^{(\lambda)}}{\partial y_i} \right| \\ &= \prod_{i=1}^n |y_i^{\lambda-1}|. \end{aligned}$$

(iii) Prior specifications

The general size and range of the transformed phenotypic data $\mathbf{y}^{(\lambda)}$ may depend strongly on λ , and so would the interpretation of the model parameters. Therefore, the prior distributions must be carefully chosen to avoid meaningless inferences. Here, we follow the approach suggested in Box & Cox (1964) to establish the conditional prior distribution, given λ , whereby the relationship between parameters is consistent under different values of λ . Calculations sketched out in Appendix A show that the prior distribution of the parameters that depend on λ is

$$\begin{aligned} & P(\exp(\mu^*) | \lambda) P(\mathbf{b} | \lambda) P(\exp(\mathbf{b}^*) | \lambda) P(\sigma_a^2 | \lambda) P(\sigma_{a^*}^2 | \lambda) P(\sigma_p^2 | \lambda) \\ & \times P(\sigma_{p^*}^2 | \lambda) P(\mu | \lambda) P(\lambda) \propto |l_\lambda|^{(2\nu - h_1 - 3)} \\ & \times (\sigma_a^2)^{-\nu/2-1} \exp \left\{ -\frac{l_\lambda^2 \times \nu S_{\sigma_a^2}}{2\sigma_a^2} \right\} \times (\sigma_{a^*}^2)^{-\nu/2-1} \\ & \times \exp \left\{ -\frac{\nu S_{\sigma_{a^*}^2}}{2\sigma_{a^*}^2} \right\} \\ & \times (\sigma_p^2)^{-\nu/2-1} \exp \left\{ -\frac{l_\lambda^2 \times \nu S_{\sigma_p^2}}{2\sigma_p^2} \right\} (\sigma_{p^*}^2)^{-\nu/2-1} \\ & \times \exp \left\{ -\frac{\nu S_{\sigma_{p^*}^2}}{2\sigma_{p^*}^2} \right\} = P_\lambda. \end{aligned} \quad (6)$$

(iv) Posterior distribution

Including in (6) the prior distribution of the parameters conditionally independent of λ , $P(\mathbf{a}, \mathbf{a}^* | \sigma_a^2, \sigma_{a^*}^2, \rho) P(\mathbf{p} | \sigma_p^2) P(\mathbf{p}^* | \sigma_{p^*}^2)$, leads to the joint prior

distribution for all model parameters. This takes the form

$$\begin{aligned}
 &P(\mathbf{a}, \mathbf{a}^*, \mathbf{p}, \mathbf{p}^*, \mu, \exp(\mu^*), \mathbf{b}, \exp(\mathbf{b}^*), \sigma_a^2, \sigma_{a^*}^2, \rho, \sigma_p^2, \sigma_{p^*}^2, \lambda) \\
 &\propto \frac{1}{(2\pi)^{n_1/2} |\mathbf{G} \otimes \mathbf{A}|^{1/2}} \exp\left\{-\frac{1}{2} \begin{pmatrix} \mathbf{a} \\ \mathbf{a}^* \end{pmatrix}' \mathbf{G}^{-1} \otimes \mathbf{A}^{-1} \begin{pmatrix} \mathbf{a} \\ \mathbf{a}^* \end{pmatrix}\right\} \\
 &\times \left(\frac{1}{\sigma_p^2 \sigma_{p^*}^2}\right)^{n_2/2} \exp\left\{-\frac{\mathbf{p}'\mathbf{p}}{2\sigma_p^2} - \frac{\mathbf{p}^*\mathbf{p}^*}{2\sigma_{p^*}^2}\right\} \times P_\lambda. \quad (7)
 \end{aligned}$$

The joint posterior distribution is proportional to the product of (4) and the prior (7).

The inferential strategy in this study is based on, firstly, estimating the modal value of λ from its marginal posterior distribution. This is done fitting jointly all the parameters of the BCHET model, including the correlation coefficient ρ . The final inferences are based on a second implementation of the BCHET model, where the value of λ is fixed at its posterior mode.

(v) *Model comparison and model checking*

Models are compared using the pseudo marginal probability of data. This is a method of comparison (Gelfand, 1996) based on the relative global fits of the models. In addition, the skewness of the conditional distribution of the data under both models is investigated using posterior predictive model checking. This provides insight into the ability of the models to address a specific feature of the data that is of primary interest. In this study, one wishes to investigate whether the transformation succeeds in inducing normality at the level of the conditional distribution of the data, which is a basic premise of the model and a necessary condition for correct inferences. Details are in Appendix B.

(vi) *McMC algorithm*

The BCHET model (1) is implemented using an McMC algorithm, where the components in the model are updated sequentially. The algorithm is based on a combination of Gibbs updates, updates based on random walk proposals and updates based on Langevin–Hastings proposals. In addition, a reparameterization of the additive genetic values (\mathbf{a}, \mathbf{a}^*) to standard normal variates (γ, γ^*) was performed to improve the mixing of the algorithm.

In summary, the BCHET model (1) is implemented using the following McMC algorithm.

• **Metropolis–Hasting steps:**

1. $(\lambda, \mu, \exp(\mu^*), \mathbf{b})$ updated jointly using a random walk proposal for λ and fully conditional proposals for $(\mu, \exp(\mu^*), \mathbf{b})$.
2. $\exp(\mathbf{b}^*)$ and (γ, γ^*) updated using Langevin–Hastings proposals.

3. $(\mathbf{p}, \mathbf{p}^*)$ updated using a fully conditional proposal for \mathbf{p} and a Langevin–Hastings proposal for \mathbf{p}^* .
4. $(\sigma_a^2, \sigma_{a^*}^2)$ updated using random walk proposals (on the log scale).
5. ρ updated using a random walk proposal.

• **Gibbs Steps:** $\sigma_p^2, \sigma_{p^*}^2$.

Convergence of the algorithm was checked informally using traceplots. We report confidence intervals of Monte Carlo estimates of various posterior means in order to give an idea of the accuracy of the Monte Carlo computations.

3. Data

(i) *Simulated data*

It is difficult to gain insight about the degree of identifiability of λ in the marginal likelihood when it is fitted jointly with ρ (together with the remaining parameters of the model), because the marginal likelihood cannot be written in closed form. One can conjecture that the number of repeated records per individual plays an important role. Therefore, a small simulation study was conducted to investigate how the quality of inferences about λ from its marginal posterior distribution is affected with increasing number of records per individual. A family structure based on the pedigree of the rabbit litter size data was simulated, with 1, 2, 3, 5 and 10 records per individual. The true value of λ was set equal to 1, and of ρ equal to -0.74 , in all cases. These true values are compared with those inferred from the marginal posterior distribution of these parameters.

(ii) *Rabbit litter size data*

The data originate from a selection experiment for uterine capacity in rabbits spanning 10 generations, carried out at the Universidad Politécnica de Valencia. Uterine capacity will be referred to as litter size hereinafter. Details of the selection experiments and data can be found in Argente *et al.* (1997) and Ibáñez *et al.* (2008). The number of individuals in the pedigree was 1281, where 929 of them had litter size records (a maximum of four records per female), and the total number of records was 2996. Reproduction was organized in discrete generations.

(iii) *Pig litter size data*

The data originate from a large-scale selection experiment for the total number of piglets born per litter (referred to as litter size hereinafter). Details of the selection experiments and data descriptions are provided in Sorensen *et al.* (2000). The number of

Table 1. MC estimates of posterior means and 95% posterior interval (95% P.I.) for λ and ρ with increasing number of repeated records per individual

	True $\lambda = 1$		True $\rho = -0.74$	
	Posterior mean	95% P.I.	Posterior mean	95% P.I.
1 Record	0.89	(0.01, 1.84)	-0.45	(-1.00, 0.20)
2 Records	1.40	(-0.80, 1.97)	-0.80	(-1.00, -0.51)
3 Records	0.99	(0.45, 1.40)	-0.84	(-1.00, -0.59)
5 Records	0.78	(0.42, 1.12)	-0.66	(-0.89, -0.42)
10 Records	0.93	(0.71, 1.18)	-0.70	(-0.80, -0.37)
Rabbit data	0.82	(0.48, 1.51)	-0.74	(-0.96, -0.50)

individuals in the pedigree was 6437, where 4078 of them had phenotypic records, and the total number of litter size records was 9778.

4. Results

(i) Simulation study

The results reported for different simulated data sets were obtained from McMC output consisting of between 1 000 000 and 3 000 000 iterations. The total number of stored samples was 10 000 for each simulated data set and these were used to infer the marginal distribution of λ .

Results are shown in Table 1. The 95% posterior intervals of λ always include the true value 1 in all simulated data sets, and inferences are sharper with increasing number of records per individual. The posterior distributions are symmetric (not shown); posterior means and modes agree to two decimal figures.

Estimates of the posterior means are clearly subject to sampling variation of the data. This frequentist variation may be characterized by replication of the experiment, but this was judged to be computationally too demanding.

The simulation study showed that the posterior correlation of λ and ρ decreased with increasing number of records per individual (not shown). In the data set mimicking the rabbit structure, this posterior correlation is equal to 0.34. The results indicate that the size of the smallest data set, corresponding to the rabbit experiment, is adequate to arrive at satisfactory inferences about features of the marginal posterior distribution of λ .

(ii) Analyses of litter size in rabbits and pigs

(a) The models

The models for the rabbit and pig data are similar to those in Sorensen & Waagepetersen (2003) and Ibáñez *et al.* (2008). The scalar μ is the mean and the scalar $\exp(\mu^*)$ is the mean environmental variance. In the rabbit data, vector \mathbf{b} contains year-season effects

with 30 levels and parity order effects with four levels (first, second, third and fourth or higher parities). The vector $\exp(\mathbf{b}^*)$ contains the environment-specific scalings effects associated with year-season and parity. Vectors \mathbf{a} and \mathbf{a}^* contain additive genetic values with 1281 levels, and vectors \mathbf{p} and \mathbf{p}^* contain permanent environmental effects with 929 levels. The parameters of the scaled inverted chi-square distributions associated with σ_a^2 , $\sigma_{a^*}^2$, σ_p^2 , $\sigma_{p^*}^2$, are $\nu = 5$, $S_{\sigma_a^2} = 0.492$, $S_{\sigma_{a^*}^2} = 0.096$, $S_{\sigma_p^2} = 0.264$ and $S_{\sigma_{p^*}^2} = 0.072$. In the pig data, \mathbf{b} is a vector containing the effects of four categorical variables: parity (eight levels), season (four levels), herd (50 levels) and type of insemination (artificial or natural). The vector $\exp(\mathbf{b}^*)$ contains the environment-specific scalings effects with two categorical variables: parity (eight levels) and type of insemination (artificial or natural). Vectors \mathbf{a} and \mathbf{a}^* contain additive genetic effects with 6437 levels, and vectors \mathbf{p} and \mathbf{p}^* contain permanent environmental effects with 4078 levels. The parameters of the scaled inverted chi-square distributions associated with σ_a^2 , $\sigma_{a^*}^2$, σ_p^2 , $\sigma_{p^*}^2$, are $\nu = 5$, $S_{\sigma_a^2} = 0.972$, $S_{\sigma_{a^*}^2} = 0.054$, $S_{\sigma_p^2} = 0.36$ and $S_{\sigma_{p^*}^2} = 0.036$. The hyperparameters of the prior distributions are taken from Sorensen & Waagepetersen (2003) and Ibáñez *et al.* (2008).

(b) The mode of λ

Inference about the modal value of λ from its marginal posterior distribution was accomplished running the McMC algorithm with a chain length equal to 1 405 000 in the rabbit data and 750 000 in the pig data. In the rabbit data, the Monte Carlo estimate of the mode of λ was equal to 1.41 and the 95% posterior interval was (1.30, 1.51). In the pig data, the corresponding figures are 1.39 and (1.32, 1.47). In both cases, the posterior distribution was symmetric (the estimated posterior means agreed with the modal values to two decimal figures). In neither case do the 95% posterior intervals of λ contain the value 1. This indicates that in both species the phenotypic data are not conditionally normally distributed under $\lambda = 1$.

Table 2. MC estimates of posterior means (first row for each model), 95% posterior intervals (second row for each model) and of 95% Monte Carlo confidence intervals (third row for each model) of chosen parameters based on models M and M_λ , for the rabbit litter size data (top) and on models M and M_λ for the pig data set (bottom).

Data	Model	σ_a^2	$\sigma_{a^*}^2$	ρ	σ_p^2	$\sigma_{p^*}^2$
Rabbits	M	0.81	0.130	-0.73	0.38	0.052
		(0.48,1.22) (0.79,0.82)	(0.056,0.230) (0.131,0.142)	(-0.89,-0.50) (-0.74,-0.72)	(0.15,0.66) (0.37,0.39)	(0.025,0.099) (0.049,0.055)
	M_λ	2.59	0.056	0.28	2.86	0.042
		(1.47,4.20) (2.55,2.64)	(0.027,0.102) (0.055,0.058)	(-0.24,0.79) (0.27,0.30)	(1.53,4.22) (2.82,2.89)	(0.02,0.084) (0.042,0.047)
Pigs	M	1.63	0.071	-0.64	0.52	0.021
		(1.24,2.05) (1.60,1.65)	(0.038,0.110) (0.069,0.073)	(-0.82,-0.45) (-0.65,-0.63)	(0.25,0.83) (0.51,0.54)	(0.010,0.038) (0.019,0.023)
	M_λ	8.17	0.037	0.70	4.15	0.017
		(5.90,10.63) (8.10,8.24)	(0.019,0.058) (0.037,0.038)	(0.44,0.98) (0.70,0.71)	(2.17,6.03) (4.10,4.20)	(0.008,0.026) (0.017,0.018)

In the rabbit data, the 95% confidence interval of the posterior mean, reflecting Monte Carlo sampling variation, was (1.405, 1.411) and (1.390, 1.398) in the pig data, indicating that the degree of accuracy of the Monte Carlo computation was satisfactory.

In order to gain insight into the sensitivity of inferences on prior assumptions, a similar analysis was conducted with scale parameters $S_{\sigma_a^2}$, $S_{\sigma_{a^*}^2}$, $S_{\sigma_p^2}$, $S_{\sigma_{p^*}^2}$ set equal to 0.124, 0.024, 0.066 and 0.018 in the rabbit data and to 0.234, 0.014, 0.090 and 0.009 in the pig data. Under these prior assumptions, the posterior mode of λ in the rabbit data is 1.40, and the 95% posterior interval for λ was (1.31, 1.51). In the pig data, the corresponding figures are 1.39 and the 95% posterior interval is (1.32, 1.46) reflecting minor differences with the original values in both cases.

(c) Variance components

Results for the rabbit data set are shown in the top rows of Table 2. The figures report Monte Carlo estimates of posterior means, 95% posterior intervals and 95% Monte Carlo confidence intervals for chosen parameters of BCHET under $\lambda=1$ (M) and BCHET under the modal value of $\lambda=1.41$ (M_λ). Estimates of σ_a^2 and σ_p^2 increase by factors of 3 and 7, respectively, when the data are analysed under M_λ , as expected, due to dependence on λ . Despite the lack of prior dependence of $\sigma_{a^*}^2$ and $\sigma_{p^*}^2$ on λ , the posterior mean of $\sigma_{a^*}^2$ decreases by a factor >2 under M_λ , while $\sigma_{p^*}^2$ is slightly reduced. The posterior mean of ρ changes from -0.73 under M to 0.28 under M_λ , and the 95% posterior interval of ρ includes zero under M_λ .

Under M, the smallest posterior mean of the environmental variance (95% posterior intervals in brackets), corresponding to year-season 30 and parity 1, was 3.28 (1.55, 5.34). The corresponding largest

number, for year-season 15 and parity 3 was 8.23 (4.61, 11.92). Under M_λ , the corresponding figures are 11.16 (5.72, 17.56) and 36.92 (26.39, 52.99).

Under M, there is one heritability for each combination of environment-specific scalings effects $\exp(\tilde{\mathbf{X}}\mathbf{b}^*)_i$ at the level of the environmental variance. The average heritability over all combinations of environmental effects is 0.12 with a minimum of 0.09 (0.04, 0.14) (year-season effect 15 and parity 3) and a maximum of 0.18 (0.09, 0.28) (year-season effect 30 and parity 1). Under M_λ , the average heritability is 0.10 with a minimum of 0.07 (0.03, 0.11) (year-season effect 15 and parity 3) and a maximum of 0.19 (0.08, 0.31) (year-season effect 30 and parity 1), reflecting little change due to the transformation. It can be shown that the prior heritability is not affected by the transformation (to the degree of approximation (A1)), a result that agrees with the conclusion from Solomon & Taylor (1999), who worked with a variance homogeneous, one-way classification model.

Figure 1 shows the estimated marginal posterior distributions of σ_a^2 , $\sigma_{a^*}^2$, σ_p^2 and $\sigma_{p^*}^2$ based on M_λ . The solid line superimposed on each graph represents the densities of the scaled inverted chi-square prior distributions of σ_a^2 , $\sigma_{a^*}^2$, σ_p^2 and $\sigma_{p^*}^2$. On the top of the figure, the variance parameters σ_a^2 , $\sigma_{a^*}^2$, σ_p^2 and $\sigma_{p^*}^2$ are assigned scaled inverted chi-squared distributions with degrees of freedom $\nu=5$ and scale parameters $S_{\sigma_a^2}=0.492$, $S_{\sigma_{a^*}^2}=0.096$, $S_{\sigma_p^2}=0.264$ and $S_{\sigma_{p^*}^2}=0.072$, resulting in prior modes equal to 0.35, 0.07, 0.19 and 0.05, respectively. The bottom of the figure shows corresponding distributions obtained setting $S_{\sigma_a^2}=0.124$, $S_{\sigma_{a^*}^2}=0.024$, $S_{\sigma_p^2}=0.066$ and $S_{\sigma_{p^*}^2}=0.018$, leading to prior modes equal to 0.09, 0.02, 0.05 and 0.013, respectively, a four-fold reduction. Prior input does not influence inferences about σ_a^2 and σ_p^2 but $\sigma_{a^*}^2$ and $\sigma_{p^*}^2$ are strongly affected. In fact, the posterior distribution of $\sigma_{a^*}^2$ shows little divergence from the

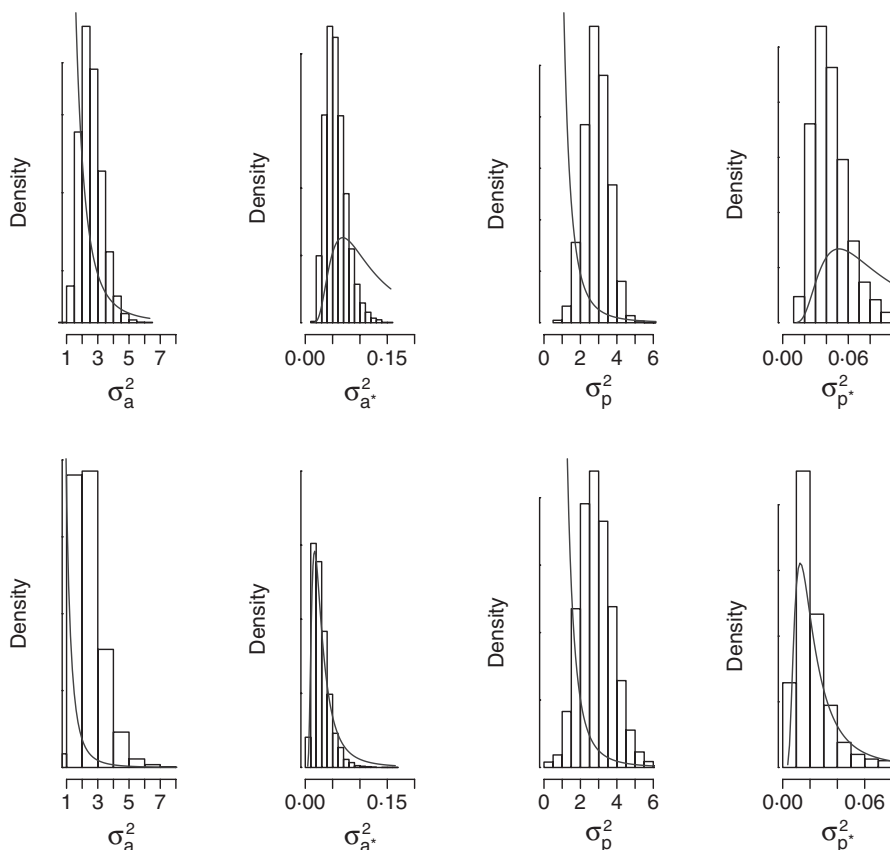


Figure 1. Top: Monte Carlo estimates of σ_a^2 and $\sigma_{a^*}^2$ (left) and σ_p^2 and $\sigma_{p^*}^2$ (right) under M_λ for rabbit data. The thick lines represent the prior scaled inverted chi-square densities with degrees of freedom $\nu=5$ and scale parameters $S_{\sigma_a^2}=0.492$, $S_{\sigma_{a^*}^2}=0.096$, $S_{\sigma_p^2}=0.264$ and $S_{\sigma_{p^*}^2}=0.072$ under M_λ . Bottom: as above, with $\nu=5$ and $S_{\sigma_a^2}=0.124$, $S_{\sigma_{a^*}^2}=0.024$, $S_{\sigma_p^2}=0.066$ and $S_{\sigma_{p^*}^2}=0.018$.

prior distribution in both scenarios, suggesting that there is little information in the data about this parameter.

In conclusion, the analysis indicates that the support for a genetically structured variance heterogeneous model in the rabbit data is markedly weaker when the data are analysed under a scale corresponding to the modal value of λ .

Estimated posterior means of dispersion parameters for the pig data set are shown in the bottom rows of Table 2. The figures show a similar pattern as for the rabbit litter size data. The posterior mean of ρ changes sign from negative under M (BCHET with $\lambda=1$) to positive under M_λ (BCHET conditional on the mode of $\lambda=1.39$). However, in contrast with the rabbit data, the 95% posterior interval of ρ does not include the value $\rho=0$. Thus, the analysis in the transformed scale results in a strong, positive correlation between additive genetic values affecting mean and variance.

Under M , the smallest posterior mean of the environmental variance, corresponding to natural insemination and parity 1, is 5.87 (5.27, 6.53). The largest posterior mean, corresponding to artificial

insemination and parity 2, is 9.33 (8.27, 10.42). Under M_λ , these figures are 33.01 (30.01, 36.02) and 58.06 (47.59, 69.62).

Under M , there is one heritability for each combination of the environment-specific scalings effects $\exp(\mathbf{Xb}^*)_i$. The average heritability is 0.16 with a minimum of 0.14 (0.11, 0.18) (artificial insemination and parity 2) and a maximum of 0.20 (0.16, 0.25) (natural insemination and parity 1). Under M_λ , the average heritability is 0.14 with a minimum of 0.12 (0.08, 0.15) (artificial insemination and parity 6) and a maximum of 0.18 (0.13, 0.23) (natural insemination and parity 1).

Figure 2 illustrates the sensitivity of inferences to prior assumptions for the pig data. In contrast with the rabbit data, a five-fold reduction in prior modes leads to very small changes in the posterior distributions, except for $\sigma_{p^*}^2$.

To give a further idea of the sensitivity of inferences to λ , the rabbit data were also analysed conditional on the lower ($\lambda=1.30$) and the upper $\lambda=1.51$ cut-off points of the 95% posterior interval for λ . The posterior distribution of $\sigma_{a^*}^2$ was hardly affected, but the posterior distribution of ρ shifted to the right with

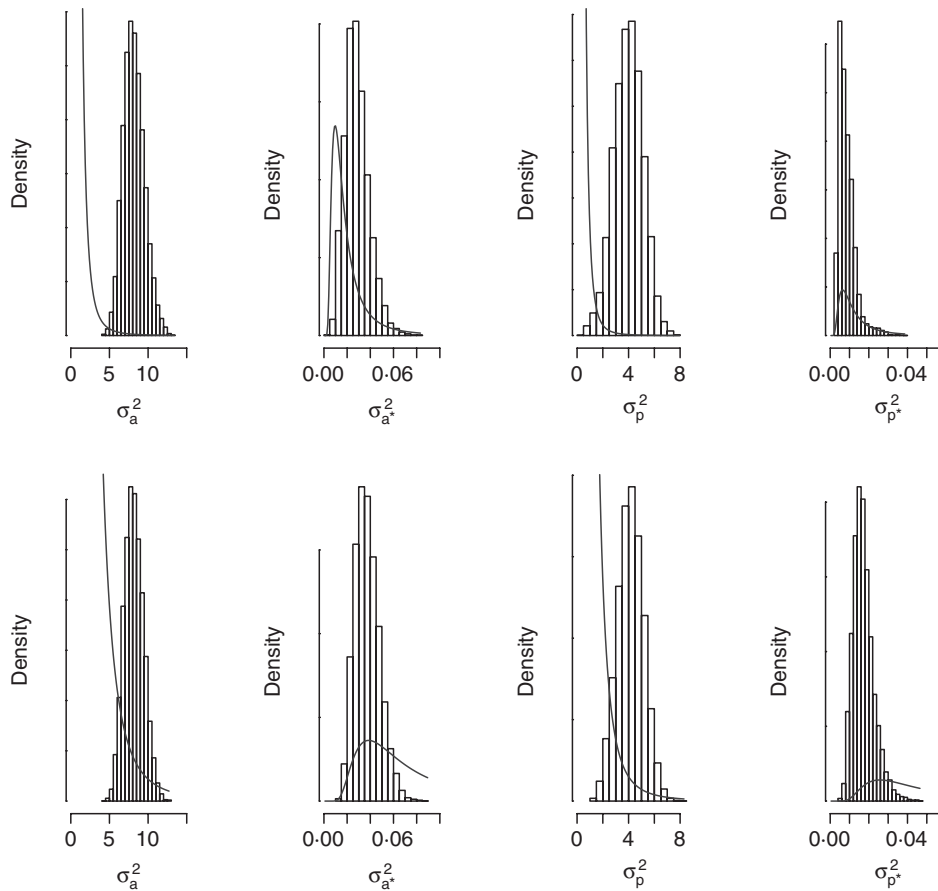


Figure 2. Top: Monte Carlo estimates of σ_a^2 and σ_a^{2*} (left) and σ_p^2 and σ_p^{2*} (right) under M_λ for pig data. The thick lines represent the prior scaled inverted χ^2 densities with degrees of freedom $\nu=5$ and scale parameters $S_{\sigma_a^2}=0.972$, $S_{\sigma_a^{2*}}=0.054$, $S_{\sigma_p^2}=0.36$ and $S_{\sigma_p^{2*}}=0.036$ under M_λ . Bottom: as above, with $\nu=5$ and scale parameters $S_{\sigma_a^2}=0.243$, $S_{\sigma_a^{2*}}=0.0135$, $S_{\sigma_p^2}=0.09$ and $S_{\sigma_p^{2*}}=0.009$.

increasing values of λ . The posterior means and 95% posterior intervals for $\lambda=1.30, 1.41$ and 1.51 were -0.10 ($-0.55, 0.37$), 0.28 ($-0.24, 0.79$) and 0.56 ($0.13, 0.96$). Although the distributions are highly overlapping, the support at $\lambda=1.51$ does not contain the value of $\rho=0$. At the modal value $\lambda=1.41$, the residual skewness is practically zero, and the assumption of conditional normality is satisfied. Lower and higher values of λ generate slight negative and positive skewness, respectively (results not shown), and this results in the observed pattern of ρ .

The 95% confidence intervals for the MC estimates of posterior means indicate that the lengths of the chains (sample sizes) result in adequate accuracy in both analyses.

(d) *Model comparison and posterior model checking*

The Monte Carlo estimate of the pseudo marginal probabilities of the rabbit data under M and under M_λ are equal to -3930.7 and -3919.5 , respectively. The corresponding figures for the pig data are -23999.0 for M and -15269.1 for M_λ . In both

cases, the best fits correspond to the analysis on the transformed scale.

Figure 3 shows histograms of posterior realizations of the discrepancy measure (17) under M and under M_λ in the rabbit data and M and M_λ in the pig data, designed to test skewness at the level of the conditional distribution of the data. In both species, the distribution of (17) discloses negative skewness when the data are analysed in the original scale (first and third histograms from left), whereas under the transformed model, the mean residual skewness is zero (second and fourth histograms from left). This indicates that the transformation was successful in inducing normality at the level of the conditional distribution of the data.

Since estimates of posterior means of σ_a^{2*} and σ_p^{2*} in the rabbit data were fairly small under M_λ , and sensitive to prior specifications, we investigated further whether heterogeneity could be accounted solely by systematic effects, without inclusion of random effects at the level of the environmental variance. In this model, labelled HM_λ , $\sigma_a^{2*}=0$ and $\sigma_p^{2*}=0$. The $\sum_{i=1}^n \log(CPO_i)$ for HM_λ is -3927.3 indicating a poorer

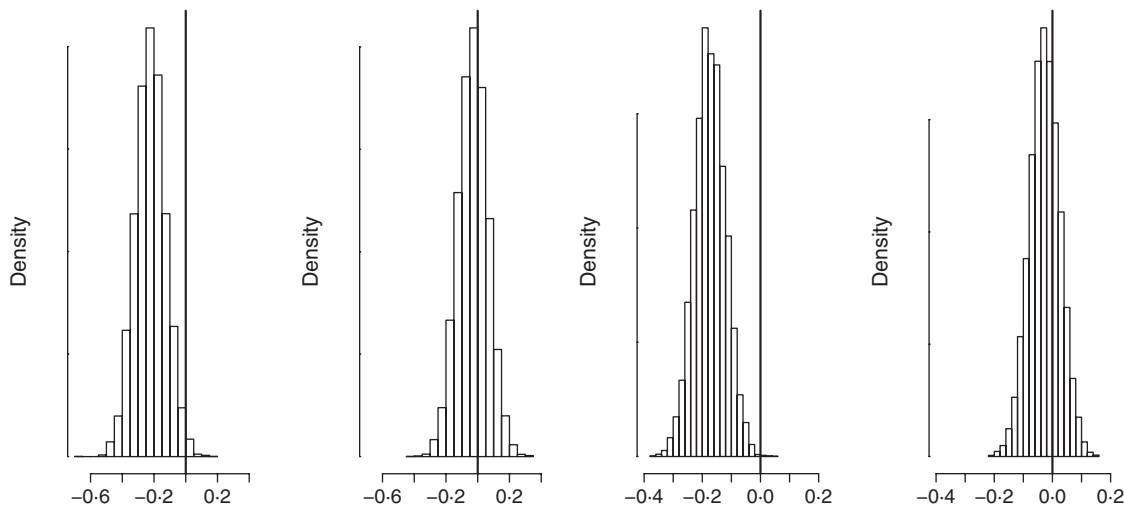


Figure 3. Histograms of posterior predictive realization of $T(\mathbf{z}^{\lambda=\lambda_0}, \theta) - T(\mathbf{z}_{rep}, \theta)$, designed to test residual skewness for the rabbit data (first two from left) and pig data (last two from left). The first and third figures correspond to analyses on the original scale; the second and fourth to analyses on the transformed scale.

fit than the full model and thus support for $\sigma_{a^*}^2$ and $\sigma_{p^*}^2$. The corresponding figure for the pig data for HM_λ is $-15\,297.1$, indicating also support for the full model.

Both data sets were also analysed using the classical infinitesimal model with homogeneous variance, conditional on the modal value of the Box–Cox parameter. The posterior mode of λ (95% posterior interval in brackets) is 1.41 (1.33, 1.49), in the rabbit data and 1.26 (1.21, 1.31) in the pig data. The $\sum_{i=1}^n \log(CPO_i)$ under the classical infinitesimal model is -4009.5 in the rabbit data and $-15\,337.23$ in the pig data. In the rabbit data, the result indicates that the global fit of the infinitesimal model is worse than that obtained fitting the genetically structured variance heterogeneity model M_λ (and also worse than model M). In the pig data, the global fit from the classical infinitesimal model is worse than that obtained fitting the genetically structured variance heterogeneity model M_λ , but better than the M model.

The overall conclusion from the analyses of both data sets is that the transformed model induces conditional normality of the sampling distribution of the data (as opposed to the untransformed model), and that the transformed model is best supported in terms of the overall fit relative to the untransformed model. Under the transformed model there is considerably weaker evidence in the rabbit data for a genetically structured variance heterogeneity model than that provided by the analysis in the original scale. In the pig data set, the transformed model provides evidence for the presence of genetic variation at the level of the environmental variance, but in contrast with the results of the analysis in the original scale, the correlation ρ between additive genetic values affecting mean and variance is strong and positive.

5. Discussion

This work has confirmed that inferences at the level of the residual variance are strongly affected by the scale on which the data are analysed. This is certainly not a surprising result, but it rather reflects the difficulties in learning about the state of nature via a statistical analysis of data using models that build on strong distributional assumptions. One option is to avoid introducing such assumptions and complex modelling, and instead use simple functions of the data specifically designed to uncover the structure associated with a genetic component at the level of variance. Uncertainty concerning the final result could be incorporated using resampling methods, such as permutation tests or the bootstrap. A limitation of this option is the difficulty of finding a measure that is free from artefacts and that signals unambiguously the existence of genetic variance on variance. Another limitation is whether the structure in the data related to the existence of genetic variance on variance is strong enough to be detected by these means. For example, we have attempted to regress sampling variance of repeated measurements within individuals in daughters on mothers and could not find a convincing pattern. A second option, followed in this work, is to use a fully parametric model that addresses specifically the putative presence of a genetic component affecting environmental variance. Support for the model is sought by comparing the quality of its fit with that of other models that neglect the presence of a genetic component on the environmental variance, using various criteria. An unambiguous favourable comparison is interpreted as evidence for the presence of a genetic component on the environmental variance. Spurious results can never be excluded using

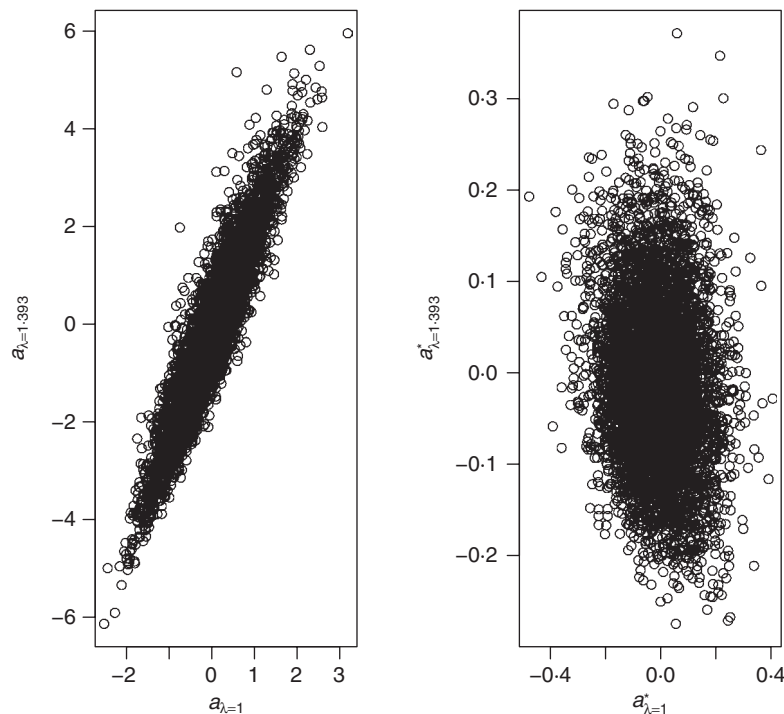


Figure 4. Left: plot of the posterior means of the vector \mathbf{a} under $M(\mathbf{a}_{\lambda=1})$ versus the posterior means of the vector \mathbf{a} under $M_{\lambda}(\mathbf{a}_{\lambda=1.393})$. Right: plot of the posterior means of the vector \mathbf{a}^* under $M(\mathbf{a}_{\lambda=1}^*)$ versus the posterior means of the vector \mathbf{a}^* under $M_{\lambda}(\mathbf{a}_{\lambda=1.393}^*)$ in pig litter size data.

this approach either. The exercise reflects merely that, among the models considered, there is one that does better than the rest. It does not provide decisive evidence.

The results of the study differ in the rabbit and pig data. The transformed rabbit litter size data yields a Monte Carlo estimate of the posterior mean for the correlation (95% posterior interval in brackets) of 0.28 (−0.24, 0.79), compared to a value equal to −0.73 (−0.89, −0.50) obtained in the untransformed data. Further, the additive genetic variance at the level of the environmental variance based on the transformed data falls by a factor of 2.4, and its posterior distribution is considerably affected by prior assumptions. Thus, under the transformed scale, the support for a genetic component at the level of the environmental variance is markedly weaker. The picture is different for the pig litter size data. The correlation changes from −0.64 (−0.82, −0.45) to 0.70 (0.44, 0.98) under the transformed model, the additive genetic variance at the level of the environmental variance falls by a factor of <2, and its posterior distribution is very little affected by prior input. On the basis of these results, the statistical evidence in favour of the genetically heterogeneous variance model is not disputed, but in the transformed model the correlation changes sign and its posterior distribution has positive support.

There has been some controversy concerning how inferences should be drawn using the Box–Cox family

of transformations. Bickel & Doksum (1981) show that joint inferences of the transformation and the remaining parameters of a linear model lead to highly correlated estimates and the marginal variances of the latter are much larger compared to the conditional variances for fixed values of the transformation parameter. Hinkley & Runger (1984) and Box & Cox (1982) argue that inferences should be drawn conditional on a fixed value of the transformation parameter, since linear model parameters have meaning only with reference to a particular scale. This is an important point that argues against reporting inferences marginalizing over the distribution λ . The consequences on statements of posterior uncertainty fitting the model conditional on λ do not change the conclusions of our analysis. For example, for the rabbit litter size data, the 95% posterior interval of ρ changes from (−0.24, 0.79) for fixed λ , to (−0.36, 0.80) when the analysis is done marginalizing over λ . In the pig data, these figures are (0.44, 0.98) and (0.36, 0.99). In the context of prediction of future data, as in De Oliveira *et al.* (1997, 2002) and Lee *et al.* (2005), it may be sensible that posterior intervals of future data incorporate uncertainty about λ .

The problem of scale has clearly important implications in an animal breeding context and we consider two scenarios. First, assume that one is interested in selecting at the level of the mean only (via predicted additive genetic values) and that the ‘correct’ model is the transformed BCHET model.

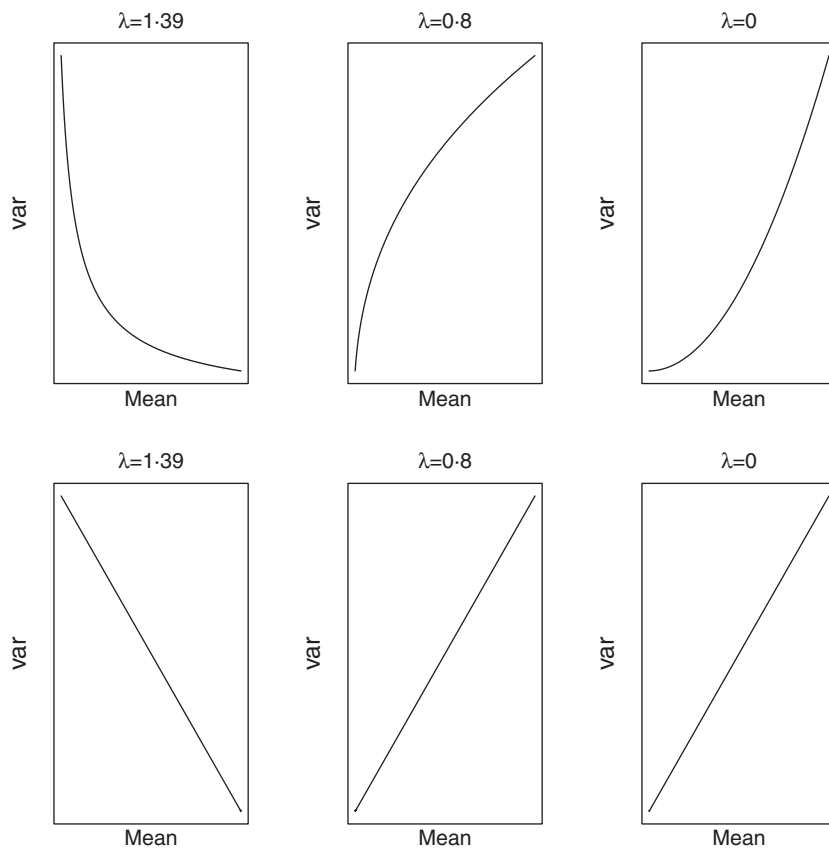


Figure 5. Relationship between variance and mean (expressions (8) and (9)) when only a_i changes and a_i^* is kept constant (top) and when a_i^* changes and a_i is kept constant (bottom), for three values of λ .

Here, the term ‘correct’ model is to be understood as the model under which the assumptions for the data in question hold. Does it make any difference whether predictions and subsequent selection decisions are obtained via the transformed BCHET or the untransformed HET model? For the pig data (the rabbit data show an almost identical pattern), the left-hand side of Figure 4 shows that there is a very close positive association between the posterior means of additive genetic values affecting the mean under both models, implying that approximately the same individuals are selected when ranked on the basis of either model. On the other hand, the right-hand side of Figure 4 discloses no association between additive genetic values affecting the variance under both models. Therefore, the efficiency of selection at the level of the variance will be impaired if selection operates on the basis of the ‘incorrect’ model.

In the second scenario, assume again that the ‘correct’ model is BCHET and selection under this model (to change the mean, the variance, or both) leads to displacements at the level of additive genetic values affecting mean and variance. One may wish to know how changes in the transformed scale translate into changes in the original scale and report results in that metric. Under the BCHET model, genetic values operating on mean and variance are additive and the

correlation between them has a clear interpretation. When transformed back to the original scale, genetic values are no longer additive and the correlation is difficult to define. However, the problem can be investigated by deriving the mean and variance of the conditional distribution of the data in the original scale, given knowledge of the transformed BCHET model. More specifically, consider the simplified version of the transformed BCHET model $\mathbf{y}^{(\lambda)}|\lambda, \mu, \mu^*, \mathbf{a}, \mathbf{a}^* \sim N(1\mu + \mathbf{Za}, \exp(1\mu^* + \mathbf{Za}^*))$. A Taylor expansion shows that the mean and variance in the original scale can be approximated by

$$E(y_i) \simeq (1 + \lambda(\mu + a_i))^{1/\lambda} + \frac{1}{2} \exp(\mu^* + a_i^*) \times (1 - \lambda)(1 + \lambda(\mu + a_i))^{1/\lambda - 2} \tag{8}$$

and

$$Var(y_i) \simeq (1 + \lambda(\mu + a_i))^{2/\lambda - 2} \exp(\mu^* + a_i^*), \tag{9}$$

respectively. For a given value of λ (and μ) one can study how the mean and variance in the original scale change due to displacements in $(\mathbf{a}, \mathbf{a}^*|\rho)$ on the transformed scale. Clearly, $(\mathbf{a}, \mathbf{a}^*|\rho)$ is non-linearly related to both (8) and (9) and Figure 5 illustrates some special cases. When $\lambda=1.39$, the variance in the original scale falls at a decreasing rate as a_i increases and a_i^* is

kept constant (implying $\rho=0$ in the transformed scale). The same scenario at a value of $\lambda=0.8/0.0$ shows that the variance in the original scale grows at a decreasing/increasing rate as a_i increases. On the other hand, selection for a_i^* in the transformed scale, keeping a_i constant (implying again $\rho=0$ in the transformed scale), leads to a negative linear relationship between mean and variance in the original scale when $\lambda=1.39$. When $\lambda=0.8$ or $\lambda=0.0$, this relationship is linear and positive, respectively. When a_i and a_i^* change simultaneously (results not shown), the relationship between mean and variance in the original scale is again highly dependent on λ and ρ . The overall conclusion is that it is not possible to make general statements about how changes taking place on the transformed scale translate into changes in the original metric.

In this work, the Box–Cox transformation was used to achieve normality (and not homogeneity of variance) at the level of the conditional distribution of the sampling model. The work can be extended in an obvious way by applying an unknown transformation to the additive genetic effects to induce normality at this level. For example under the homogeneous variance model, if data are conditionally normally distributed, but additive genetic effects a are not, then the marginal distribution of the data may display skewness. This can lead to spurious ρ when the (incorrect) HET model is fitted to the data. Other situations that can lead to misleading inferences include the presence of unaccounted mixtures, the wrong functional relationship between mean and variance, the wrong model for the sampling distribution of the data, or the presence of an unaccounted correlation between additive genetic and environmental effects. This makes it strikingly clear that an attempt to understand the state of nature via a statistical analysis of data must be regarded only as a first step until more fundamental knowledge becomes available.

Appendix A

The starting point in the derivation of the prior distribution is to assume that it admits the factorization

$$\begin{aligned}
 &P(\mu, \exp(\mu^*), \mathbf{b}, \exp(\mathbf{b}^*), \mathbf{a}, \mathbf{a}^*, \mathbf{p}, \mathbf{p}^*, \sigma_a^2, \sigma_a^{*2}, \rho, \sigma_p^2, \sigma_p^{*2}, \lambda) \\
 &= P(\mathbf{a}, \mathbf{a}^*, \mathbf{p}, \mathbf{p}^* | \sigma_a^2, \sigma_a^{*2}, \rho, \sigma_p^2, \sigma_p^{*2}) \\
 &P(\mu, \exp(\mu^*), \mathbf{b}, \exp(\mathbf{b}^*), \sigma_a^2, \sigma_a^{*2}, \rho, \sigma_p^2, \sigma_p^{*2} | \lambda) P(\lambda) \\
 &= P(\mathbf{a}, \mathbf{a}^* | \sigma_a^2, \sigma_a^{*2}, \rho) P(\mathbf{p} | \sigma_p^2) P(\mathbf{p}^* | \sigma_p^{*2}) P(\sigma_a^2 | \lambda) P(\sigma_a^{*2} | \lambda) \\
 &\times P(\rho | \lambda) P(\sigma_p^2 | \lambda) P(\sigma_p^{*2} | \lambda) \\
 &\times P(\mu | \lambda) P(\exp(\mu^*) | \lambda) P(\mathbf{b} | \lambda) P(\exp(\mathbf{b}^*) | \lambda) P(\lambda). \tag{A1}
 \end{aligned}$$

Let $(\mu_\lambda, \exp(\mu_\lambda^*), \mathbf{b}_\lambda, \exp(\mathbf{b}_\lambda^*), \sigma_{a,\lambda}^2, \sigma_{a,\lambda}^{*2}, \sigma_{p,\lambda}^2, \sigma_{p,\lambda}^{*2}, \rho_\lambda)$ be the parameters for a given λ , and hence, for $\lambda=1$, the parameters are represented by $(\mu_{\lambda=1}, \exp(\mu_{\lambda=1}^*), \mathbf{b}_{\lambda=1},$

$\exp(\mathbf{b}_{\lambda=1}^*), \sigma_{a,\lambda=1}^2, \sigma_{p,\lambda=1}^2, \sigma_{a^*,\lambda=1}^2, \sigma_{p^*,\lambda=1}^2)$. The relationship between these parameters under an arbitrary λ and $\lambda=1$ are derived as follows, assuming that the relationship between the transformed datum $y_i^{(\lambda)}$ and the untransformed datum $y_i^{(\lambda=1)}$ is approximately (Box & Cox, 1964)

$$y_i^{(\lambda)} \simeq k_\lambda + l_\lambda y_i^{(\lambda=1)}, \tag{A2}$$

where

$$\begin{aligned}
 l_\lambda &= \{J(\lambda; \mathbf{y})\}^{1/n} = \left\{ \prod_{i=1}^n \left| \frac{\partial y_i^{(\lambda)}}{\partial y_i} \right| \right\}^{1/n}, \\
 k_\lambda &= \frac{(\lambda - 1)\{J(\lambda, \mathbf{y})\}^{1/n} - 1}{\lambda}.
 \end{aligned}$$

The model for the untransformed datum can be formulated as

$$\begin{aligned}
 y_i^{(\lambda=1)} &= \mu_{\lambda=1} + \mathbf{x}'_i \mathbf{b}_{\lambda=1} + \sigma_{a,\lambda=1} N(0, 1) + \sigma_{p,\lambda=1} N(0, 1) \\
 &+ \exp(\mu_{\lambda=1}^*)^{1/2} \exp(\tilde{\mathbf{X}}'_i \mathbf{b}_{\lambda=1}^*)^{1/2} \\
 &\times \exp(\sigma_{a^*,\lambda=1} + \sigma_{p^*,\lambda=1})^{1/2} N(0, 1), \tag{A3}
 \end{aligned}$$

where \mathbf{X}'_i is an $(1 \times h_1)$ row vector associated with i th observation from the matrix \mathbf{X} with the rank h_1 , and $\tilde{\mathbf{X}}'_i$ is an $(1 \times h_2)$ row vector associated with i th observation from the matrix $\tilde{\mathbf{X}}$ with the rank h_2 . Substituting (A3) in (A1),

$$\begin{aligned}
 y_i^{(\lambda)} &\simeq k_\lambda + l_\lambda \mu_{\lambda=1} + l_\lambda (\mathbf{x}'_i \mathbf{b}_{\lambda=1}) + l_\lambda \sigma_{a,\lambda=1} N(0, 1) \\
 &+ l_\lambda \sigma_{p,\lambda=1} N(0, 1) + l_\lambda \exp(\mu_{\lambda=1}^*)^{1/2} \exp(\tilde{\mathbf{X}}'_i \mathbf{b}_{\lambda=1}^*)^{1/2} \\
 &\times \exp(\sigma_{a^*,\lambda=1} + \sigma_{p^*,\lambda=1})^{1/2} N(0, 1), \tag{A4}
 \end{aligned}$$

which shows that to the level of approximation (A2), the relationship between the parameters of the transformed and untransformed models is

$$\begin{aligned}
 \mu_\lambda &\simeq k_\lambda + l_\lambda \mu_{\lambda=1}, \quad \mathbf{b}_\lambda \simeq l_\lambda \mathbf{b}_{\lambda=1}, \\
 \exp(\mu_\lambda^*) &\simeq l_\lambda^2 \exp(\mu_{\lambda=1}^*), \quad \exp(\mathbf{b}_\lambda^*) \simeq \exp(\mathbf{b}_{\lambda=1}^*), \\
 \sigma_{a,\lambda}^2 &\simeq l_\lambda^2 \sigma_{a,\lambda=1}^2, \quad \sigma_{p,\lambda}^2 \simeq l_\lambda^2 \sigma_{p,\lambda=1}^2, \quad \rho_\lambda \simeq \rho_{\lambda=1} \\
 \sigma_{a^*,\lambda}^2 &\simeq \sigma_{a^*,\lambda=1}^2, \quad \sigma_{p^*,\lambda}^2 \simeq \sigma_{p^*,\lambda=1}^2. \tag{A5}
 \end{aligned}$$

These relationships are used together with the transformation theorem to construct a conditional prior distribution for an arbitrary λ based on a specified prior distribution for $\lambda=1$. The prior distribution assumptions under $\lambda=1$ are as follows. Improper uniform distributions are assigned to $\mu_{\lambda=1}, \exp(\mu_{\lambda=1}^*), \mathbf{b}_{\lambda=1}, \exp(\mathbf{b}_{\lambda=1}^*)$; scaled inverted chi-squared distributions with degrees of freedom ν and scale parameters $S_{\sigma_a^2}, S_{\sigma_a^{*2}}, S_{\sigma_p^2}, S_{\sigma_p^{*2}}$ are assigned to the variance parameters $\sigma_{a,\lambda=1}^2, \sigma_{a^*,\lambda=1}^2, \sigma_{p,\lambda=1}^2, \sigma_{p^*,\lambda=1}^2$; the parameter $\rho_{\lambda=1}$ is assigned a uniform prior bounded in $(-1, 1)$ and the parameter λ is assigned a uniform prior bounded in $(-3, 3)$.

Appendix B

(i) *Pseudo marginal probability of the data*

The pseudo log-marginal probability of the data is a measure of model comparison (Gelfand, 1996) and is defined and computed as follows. Consider data vector $\mathbf{y}' = (y_i, \mathbf{y}'_{-i})$, where y_i is the i th datum, and \mathbf{y}_{-i} is the vector of data with the i th datum deleted. For a particular model, the conditional predictive distribution can be written as

$$P(y_i | \mathbf{y}_{-i}) = \int P(y_i | \theta, \mathbf{y}_{-i}) P(\theta | \mathbf{y}_{-i}) d\theta, \tag{B1}$$

and can be interpreted as the likelihood of each datum given the remainder of the data, where θ is the vector of all model parameters. The actual value of $P(y_i | \mathbf{y}_{-i})$ is known as the conditional predictive ordinate (CPO) for the i th observation. A Monte Carlo approximation of the CPO in (B1) for observation i is given by

$$P(y_i | \mathbf{y}_{-i}) = \left[\frac{1}{N} \sum_{j=1}^N \frac{1}{P(y_i | \theta^{(j)})} \right]^{-1}, \tag{B2}$$

where N is the number of MCMC draws and $\theta^{(j)}$ is the j th draw from the posterior distribution of θ . In this study, from (4), the sampling distribution of the untransformed datum y_i has density

$$P(y_i | \theta) = P(y_i^{(\lambda)} | \theta^{(j)}, \lambda) J(y_i, \lambda),$$

In (B2), $P(y_i | \theta^{(j)})$ therefore takes the form

$$P(y_i^{(\lambda)} | \theta^{(j)}, \lambda) | y_i^{\lambda-1}.$$

A common summary statistic based on the CPO is $\sum_{i=1}^n \log(\text{CPO}_i)$, known as the pseudo log marginal probability of the data. A larger value of $\sum_{i=1}^n \log(\text{CPO}_i)$ indicates a better relative fit of a model.

(ii) *Posterior predictive model checking*

A technique for checking the fit of a model to observed data \mathbf{y} is to draw simulated values \mathbf{y}_{rep} from the posterior predictive distributions of replicated data and compare \mathbf{y}_{rep} with the observed data (Gelman *et al.*, 2004). Any systematic differences between the observed and the simulated data indicate potential failings of the model. More specifically, the idea is to define a so-called discrepancy measure $T(\mathbf{y}, \theta)$, which depends on the data and perhaps also on θ , an unknown vector of parameters of the model under scrutiny, called null-model below. In this work, the null model is the model for the untransformed data. The measure T is specifically designed to test a particular feature of the data y that may be of scientific relevance. Replicated data are then simulated from the posterior predictive distribution of the null model,

given $\theta^{(j)}$ for the j th iteration, from which $T(\mathbf{y}_{\text{rep}}, \theta^{(j)})$ is constructed and compared with $T(\mathbf{y}_{\text{rep}}, \theta^{(j)})$. Differences between the T s may be due to sampling or due to the inability of the null model to account for the feature of the observed data disclosed by the discrepancy measure T .

In this study, we are concerned with investigating the residual skewness under the model with $\lambda = \lambda_0$. This is accomplished using the discrepancy measure

$$T(\mathbf{z}^{(\lambda=\lambda_0)}, \theta^{(j)}) = \frac{\frac{1}{n} \sum_{i=1}^n (z_i^{(\lambda=\lambda_0)})^3}{\left(\frac{1}{n} \sum_{i=1}^n (z_i^{(\lambda=\lambda_0)})^2 - \left(\frac{1}{n} \sum_{i=1}^n z_i^{(\lambda=\lambda_0)} \right)^2 \right)^{3/2}}, \tag{B3}$$

where

$$z_i^{(\lambda=\lambda_0)} = \frac{(y_i^{(\lambda=\lambda_0)} - (1\mu^{(j)} + \mathbf{Xb}^{(j)} + \mathbf{Za}^{(j)} + \mathbf{Wp}^{(j)}))_i}{\exp(1\mu^* + \tilde{\mathbf{Xb}}^{*(j)} + \mathbf{Za}^{*(j)} + \mathbf{Wp}^{*(j)})_i^{1/2}}.$$

and $\theta' = (\mu, \exp(\mu^*), \mathbf{b}', \exp(\mathbf{b}^*)', \mathbf{a}', \mathbf{a}^*, \sigma_a^2, \sigma_a^{*2}, \rho, \mathbf{p}', \mathbf{p}^*, \sigma_p^2, \sigma_p^{*2})$. The simulated datum $\mathbf{y}_{\text{rep}} = (y_{\text{rep},i})_{i=1}^n$ follows $N(1\mu + \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp}, \mathbf{R})$, that is

$$z_{\text{rep},i} = \frac{(y_{\text{rep},i} - (1\mu + \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp}))_i}{\exp(1\mu^* + \tilde{\mathbf{Xb}}^* + \mathbf{Za}^* + \mathbf{Wp}^*)_i^{1/2}} \sim N(0, 1).$$

Therefore $\mathbf{z}_{\text{rep}} = (z_{\text{rep},i})_{i=1}^n$ is directly simulated from a standard normal distribution, followed by the computation of $T(\mathbf{z}_{\text{rep}}, \theta^{(j)})$ and $T(\mathbf{z}^{(\lambda=\lambda_0)}, \theta^{(j)})$ in each MCMC iteration. Depending on whether the null model or the transformed model is under scrutiny, $\lambda_0 = 1$ or λ_0 is equal to its posterior mode. A graphical display involves plotting histograms of $T(\mathbf{z}^{(\lambda=\lambda_0)}, \theta) - T(\mathbf{z}_{\text{rep}}, \theta)$ for both models and for the rabbit and pig litter size data.

References

Ansel, J., Bottin, H., Rodriguez-Beltran, C., Damon, C., Nagarajan, M., Fehrmann, S., François, J. & Yvert, G. (2008). Cell-to-cell stochastic variation in gene expression is a complex genetic trait. *PLoS Genetics* **4**, doi:10.1371/journal.pgen.1000049.

Argente, M., Santacreu, M., Climent, A., Bolet, G. & Blasco, A. (1997). Divergent selection for Uterine Capacity in Rabbits. *Journal of Animal Science* **75**, 2350–2354.

Bickel, P. J. & Doksum, K. A. (1981). An analysis of transformations revisited. *Journal of the American Statistical Association* **76**, 296–311.

Box, G. E. P. & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B* **26**, 211–252.

Box, G. E. P. & Cox, D. R. (1982). An analysis of transformations revisited, rebutted. *Journal of the American Statistical Association* **77**, 209–210.

De Oliveira, V., Kedem, B. & Short, D. (1997). Bayesian prediction of transformed Gaussian fields. *Journal of the American Statistical Association* **92**, 1422–1433.

- De Oliveira, V., Fokianos, K. & Kedem, B. (2002). Bayesian transformed Gaussian random fields. *Japanese Journal of Applied Statistics* **31**, 175–187.
- Falconer, D. S. & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Harlow, Essex, UK: Longman.
- Foulley, J. L. & Quaas, R. L. (1995). Heterogeneous variances in Gaussian linear mixed models. *Genetics, Selection, Evolution* **27**, 211–228.
- Foulley, J. L., San Cristobal, M., Gianola, D. & Im, S. (1992). Marginal likelihood and Bayesian approaches to the analysis of heterogeneous residual variances in mixed linear Gaussian models. *Computational Statistics and Data Analysis* **13**, 291–305.
- Gelfand, A. E. (1996). Model determination using sampling-based methods. In *Markov Chain Monte Carlo in Practice* (ed. W. R. Gilks, S. Richardson & D. J. Spiegelhalter), pp. 145–161. London: Chapman and Hall.
- Gelman, A., Carlin, J., Stern, H. & Rubin, D. (2004). *Bayesian Data Analysis*. London: Chapman and Hall.
- Gianola, D., Im, S., Fernando, R. L. & Foulley, J. L. (1990). Mixed model methodology and the Box-Cox theory of transformations. In *Advances in Statistical Methods for Genetic Improvement of Livestock* (ed. D. Gianola & K. Hammond), pp. 15–40. Berlin: Springer-Verlag.
- Gianola, D., Foulley, J. L., Fernando, R. L., Henderson, C. R. & Weigel, K. A. (1992). Estimation of heterogeneous variances using empirical Bayes methods: theoretical considerations. *Journal of Dairy Science* **75**, 2805–2823.
- Gutierrez, J. P., Nieto, B., Piqueras, P., Ibáñez, N. & Salgado, C. (2006). Genetic parameters for canalisation analysis of litter size and litter weight at birth in mice. *Genetics, Selection, Evolution* **38**, 445–462.
- Hill, W. G. & Zhang, X. S. (2004). Effects on phenotypic variability of directional selection arising through genetic differences in residual variability. *Genetical Research* **83**, 121–132.
- Hinkley, D. V. & Runger, G. (1984). The analysis of transformed data. *Journal of the American Statistical Association* **79**, 302–320.
- Ibáñez, N., Varona, L., Sorensen, D. & Noguera, J. L. (2007). A study of heterogeneity of environmental variance for slaughter weight in pigs. *Animal* **2**, 19–26.
- Ibáñez, N., Sorensen, D., Waagepetersen, R. & Blasco, A. (2008). Selection for environmental variation: a statistical analysis and power calculations to detect response. *Genetics* **180**, 2209–2226.
- Ibe, S. N. & Hill, W. G. (1988). Transformation of poultry egg production data to improve normality, homoscedasticity and linearity of regression. *Journal of Animal Breeding and Genetics* **105**, 231–240.
- Lee, J. C., Lin, T. I., Lee, K. J. & Hsu, Y. L. (2005). Bayesian analysis of Box–Cox transformed linear mixed models with ARMA(p,q) dependence. *Journal of Statistical Planning and Inference* **133**, 435–451.
- Mackay, T. F. C. & Lyman, R. F. (2005). *Drosophila bristles* and the nature of quantitative genetic variation. *Philosophical Transactions of the Royal Society B* **360**, 1513–1527.
- Mulder, H. A., Bijma, P. & Hill, W. G. (2007). Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics* **175**, 1895–1910.
- Mulder, H. A., Bijma, P. & Hill, W. G. (2008). Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance. *Genetics, Selection, Evolution* **40**, 37–59.
- Ordas, B., Malvar, R. A. & Hill, W. G. (2008). Genetic variation and quantitative trait loci associated developmental stability and the environmental correlation between traits in maize. *Genetical Research* **90**, 385–395.
- Ros, M., Sorensen, D., Waagepetersen, R., Dupont-Nivet, M., SanCristobal, M., Bonnet, J.-C. & Mallard, J. (2004). Evidence for genetic control of adult weight plasticity in the snail *Helix aspersa*. *Genetics* **168**, 2089–2097.
- Rowe, S., White, I., Avendano, S. & Hill, W. G. (2006). Genetic heterogeneity of residual variance in broiler chickens. *Genetics, Selection, Evolution* **38**, 617–635.
- San Cristobal, M., Foulley, J. L. & Manfredi, E. (1993). Inference about multiplicative heteroscedastic components of variance in a mixed linear Gaussian model with an application to beef cattle breeding. *Genetics, Selection, Evolution* **25**, 3–30.
- San Cristobal-Gaudy, M., Elsen, J. M., Bodin, L. & Chevalet, C. (1998). Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. *Genetics, Selection, Evolution* **30**, 423–451.
- Solomon, P. & Taylor, J. (1999). Orthogonality and transformations in variance components models. *Biometrika* **86**(2), 289–300.
- Sorensen, D. & Waagepetersen, R. (2003). Normal linear models with genetically structured residual variance heterogeneity: a case study. *Genetical Research* **82**, 207–222.
- Sorensen, D., Vernersen, A. & Andersen, S. (2000). Bayesian analysis of response to selection: a case study using litter size in Danish Yorkshire pigs. *Genetics* **156**, 283–295.
- Wolc, A., White, I., Avendano, S. & Hill, W. G. (2009). Genetic variability in residual variation in body weight and conformation scores in broiler chicken. *Poultry Science* **88**, 1156–1161.
- Zhang, X. & Hill, W. (2005). Evolution of the environmental component of phenotypic variance: stabilizing selection in changing environments and the cost of homogeneity. *Evolution* **59**, 1237–1244.