Reaction norm model with unknown environmental covariate to analyze heterosis by environment interaction

G. Su,¹ P. Madsen, and M. S. Lund

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

ABSTRACT

Crossbreeding is currently increasing in dairy cattle production. Several studies have shown an environment-dependent heterosis [i.e., an interaction between heterosis and environment ($H \times E$)]. An $H \times E$ interaction is usually estimated from a few discrete environment levels. The present study proposes a reaction norm model to describe $H \times E$ interaction, which can deal with a large number of environment levels using few parameters. In the proposed model, total heterosis consists of an environment-independent part, which is described as a function of heterozygosity, and an environment-dependent part, which is described as a function of heterozygosity and environmental value (e.g., herd-year effect). A Bayesian approach is developed to estimate the environmental covariates, the regression coefficients of the reaction norm, and other parameters of the model simultaneously in both linear and nonlinear reaction norms. In the nonlinear reaction norm model, the $H \times E$ is approximated using linear splines. The approach was tested using simulated data, which were generated using an animal model with a reaction norm for heterosis. The simulation study includes 4 scenarios (the combinations of moderate vs. low heritability and moderate vs. low herd-year variation) of $H \times E$ interaction in a nonlinear form. In all scenarios, the proposed model predicted total heterosis very well. The correlation between true heterosis and predicted heterosis was 0.98 in the scenarios with low herd-year variation and 0.99 in the scenarios with moderate herd-year variation. This suggests that the proposed model and method could be a good approach to analyze $H \times E$ interactions and predict breeding values in situations in which heterosis changes gradually and continuously over an environmental gradient. On the other hand, it was found that a model ignoring $H \times E$ interaction did not significantly harm the prediction of breeding value under the simulated scenarios in which the variance

for environment-dependent heterosis effects was small (as it generally is), and sires were randomly used over production environments.

Key words: Gibbs sampler, heterosis by environment interaction, reaction norm model, spline regression

INTRODUCTION

Crossbreeding has been widely applied in plants and animals to make use of heterosis, to relieve inbreeding, and to introduce desirable genes into target populations. Several studies have reported an environment-dependent heterosis [i.e., heterosis $\times$ environment interaction ($H \times E$)]. In general, heterosis is larger in a suboptimal environment than in an optimal environment, depending on traits and species (Barlow, 1981).

An $H \times E$ interaction is usually examined by comparing the measures of heterosis in a few distinct environments or environment clusters. In beef cattle, many previous studies have analyzed heterosis $\times$ forage (or feeding) environment interaction (Arthur et al., 1999; Brown et al., 2001; Phillips et al., 2001). In dairy cattle, Bryant et al. (2007) analyzed breed performance and heterosis in different environments including herd-average of yield (4 levels), heat load index (4 levels), herd size (3 levels), and altitude (3 levels). The traditional method is useful in situations with few levels of environments. Generally, grouping environments could have some disadvantages. First, if a continuous underlying scale exists, the grouping could be more or less arbitrary (Strandberg, 2006). Second, the grouping ignores the variation of environment effects within group. Third, it may lead to inefficient estimation of parameters.

To deal with a large number of environment levels (e.g., herd-years) in the analysis of $H \times E$ interaction, an alternative approach could be to apply reaction norm model. Reaction norm models have been used in analysis of additive genetic and environment interaction, especially in dairy cattle (Calus et al., 2002; Kolmodin et al., 2002; Oseni et al., 2004). Su et al. (2006) developed a method to estimate environmental value simultaneously with other parameters in a reaction norm model.

Received June 30, 2008.

Accepted December 3, 2008.

¹Corresponding author: guosheng.su@agrsci.dk

It is a statistical challenge to distinguish production environment effect, heterosis, and breeding value in a hybrid population, provided $H \times E$ interaction. The objectives of the present study were: a) to propose a reaction norm model to analyze $H \times E$ interaction, b) to develop a computing method for the proposed model, and c) to test the algorithm for the proposed model and method using simulated dairy cattle data. In addition, the effect of ignoring $H \times E$ interaction on predicted breeding value was investigated, based on the simulated dairy cattle data.

MATERIALS AND METHODS

Model

The proposed model partitions heterosis into 2 components. The first one is environment-independent heterosis, which is described as a regression on heterozygosity, and the second is environment-dependent heterosis, which is described as a regression on the product of heterozygosity and environmental value.

The proposed reaction norm model is an extension of a traditional regression of heterosis on heterozygosity (h) based on the dominance theory. In a traditional regression model, heterosis is expressed as heterosis = bh .

In the situation of $H \times E$ interaction, the regression coefficients (b) are different in different environments. Assume that environments can be quantified as an environmental gradient, the regression coefficients can be described as a function of environmental values:

$$b_i = \beta_0 + \sum_{k=1}^m \beta_k \phi_k(u_i),$$

where b_i is the regression coefficient and u_i is the environmental value for environment i and $\phi_k(u_i)$ is the k th function of environmental value u_i .

For example, there are 5 environments with an environmental gradient $\mathbf{u}' = (10, 15, 20, 25, 30)$, and the regression coefficients of heterosis on heterozygosity in the 5 environments are $\mathbf{b}' = (10, 12, 14, 16, 18)$. It can be seen that there is a linear relationship between $\mathbf{b}$ and $\mathbf{u}$. Thus, $\mathbf{b}$ can be described by a function:

$$\mathbf{b} = \beta_0 + \mathbf{u}\beta_1,$$

which gives solutions $\beta_0 = 6$ and $\beta_1 = 0.4$.

In the situation of nonlinear relationship between $\mathbf{b}$ and $\mathbf{u}$, say $\mathbf{b}' = (13.0, 15.5, 17.0, 17.5, 17.0)$, $\mathbf{b}$ can be described by a function as:

$$\mathbf{b} = \beta_0 + \mathbf{u}\beta_1 + \mathbf{u}_2\beta_2,$$

where $\mathbf{u}_2$ is the vector of u_i^2 . This results in $\beta_0 = 5$, $\beta_1 = 1$, and $\beta_2 = -0.02$.

As demonstrated above, in the case of a large number of environment levels, $H \times E$ interaction can be accounted for using a reaction norm with few parameters. Therefore, an observation (y_{ij}) involved in $H \times E$ interaction can be modeled as:

$$y_{ij} = \mu + u_i + \beta_0 h_j + \sum_{k=1}^m \beta_k h_j \phi_k(u_i) + a_j + e_{ij}, \quad [1]$$

where μ is the intercept, u_i is the effect of production environment i , h_j is the heterozygosity of animal j , $\phi_k(u_i)$ is the k th function of u_i , β_0 is the regression coefficient associated with h_j , β_k is the regression coefficient associated with $h_j \phi_k(u_i)$, a_j is the additive genetic effect of animal j , and e_{ij} is the random residual.

Restricting to linear $H \times E$ interaction, the model is simplified as:

$$y_{ij} = \mu + u_i + \beta_0 h_j + \beta_1 h_j u_i + a_j + e_{ij}.$$

For nonlinear $H \times E$ interaction, it is difficult to pre-determine appropriate functions to describe nonlinear pattern. Moreover, because the model includes various functions of environmental values, it is difficult to construct mixed model equations in terms of production environment effects. Therefore, the proposed model applies linear spline regression to fit the total heterosis [i.e., using linear regression for each subrange of independent variables (environmental values in the current study) to approximate a nonlinear regression]. Let T_k ($k = 1, 2, \dots, m$) be the value of knot k (the start point of subrange k). If $T_k \leq u_i < T_{k+1}$:

$$\phi_k(u_i) = \frac{T_{k+1} - u_i}{T_{k+1} - T_k},$$

and

$$\phi_{k+1}(u_i) = 1 - \phi_k(u_i).$$

Thus, each u_i contributes to 2 covariates, and $\phi_{k+1}(u_i) + \phi_k(u_i) = 1$. It indicates $h_j \phi_k(u_i) + h_j \phi_{k+1}(u_i) = h_j$, consequently, β_0 and β_k are not identifiable. Thus, the total heterosis is expressed as:

$$\begin{aligned} & \beta_0 h_j + \beta_k h_j \phi_k(u_i) + \beta_{k+1} h_j \phi_{k+1}(u_i) \\ &= \beta_0 [h_j \phi_k(u_i) + h_j \phi_{k+1}(u_i)] + \beta_k h_j \phi_k(u_i) + \beta_{k+1} h_j \phi_{k+1}(u_i) \\ &= (\beta_0 + \beta_k) h_j \phi_k(u_i) + (\beta_0 + \beta_{k+1}) h_j \phi_{k+1}(u_i). \end{aligned}$$

Let $\gamma_k = \beta_0 + \beta_k$, $\gamma_{k+1} = \beta_0 + \beta_{k+1}$, and so on (notice that $\gamma_0 = \beta_0$ and $\gamma_1 = \beta_1$ in the linear reaction norm model). The model for nonlinear $H \times E$ is written as:

$$y_{ij} = \mu + u_i + \sum_{k=1}^m \gamma_k h_j \phi_k(u_i) + a_j + e_{ij}. \quad [2]$$

The model for linear $H \times E$ is:

$$y_{ij} = \mu + u_i + \gamma_0 h_j + \gamma_1 h_j u_i + a_j + e_{ij}. \quad [3]$$

In matrix form, the reaction norm model can be written as

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_u \mathbf{u} + \mathbf{H}\boldsymbol{\gamma} + \mathbf{Z}_a \mathbf{a} + \mathbf{e}, \quad [4]$$

where $\mathbf{b}$ is a vector of the fixed effects (location parameters in Bayesian setting) except for the regressions associating with environmental value $\mathbf{u}$, $\mathbf{u}$ is a vector of production environment effect, $\boldsymbol{\gamma}$ is a vector of regressions on the product of heterozygosity and function of environment effect (reaction norm), $\mathbf{a}$ is a vector of additive genetic effect, $\mathbf{e}$ is a vector of residuals, and $\mathbf{X}$, $\mathbf{Z}_u$, $\mathbf{H}$, and $\mathbf{Z}_a$ are incidence matrixes. The nonzero elements of the matrices $\mathbf{H}$ are the functions of environmental value $\mathbf{u}$. In the model, $\mathbf{u}$ is treated as random to avoid problems with identifiability. To make description convenient, in the context, location parameters ($\mathbf{b}$, $\mathbf{u}$, $\boldsymbol{\gamma}$, and $\mathbf{a}$) are expressed as fixed and random effects (which are usually used in likelihood setting), though the terms are seldom used in Bayesian settings.

The conditional distribution of $\mathbf{y}$ is assumed to be normal having the form:

$$\mathbf{y} | \mathbf{b}, \boldsymbol{\gamma}, \mathbf{u}, \mathbf{a}, \mathbf{R} \sim N(\mathbf{X}\mathbf{b} + \mathbf{H}\boldsymbol{\gamma} + \mathbf{Z}_u \mathbf{u} + \mathbf{Z}_a \mathbf{a}, \mathbf{R}),$$

where $\mathbf{R}$ is the residual covariance matrix. To make demonstration easy, in the context, the residual variance is assumed to be homoscedastic and the residuals are independent of each other, such that $\mathbf{R} = \mathbf{I}\sigma_e^2$, where $\mathbf{I}$ is the identity matrix and σ_e^2 is the residual variance.

To estimate the environmental covariates, the regression coefficients of the reaction norm, and the other parameters of the model simultaneously, a Bayesian Gibbs sampling approach was developed on the basis of the methodology reported by Su et al. (2006).

Prior Distributions and Joint Posterior Distribution

The prior distributions of $\mathbf{b}$ and $\boldsymbol{\gamma}$ are assumed to be improper uniform, $p(\mathbf{b}) \propto \text{constant}$, $p(\boldsymbol{\gamma}) \propto \text{constant}$.

The random vectors of $\mathbf{u}$ and $\mathbf{a}$ are assumed to have normal prior distributions:

$$\mathbf{u} \sim N(\mathbf{0}, \mathbf{I}\sigma_u^2)$$

$$\mathbf{a} \sim N(\mathbf{0}, \mathbf{A}\sigma_a^2),$$

where σ_u^2 is the variance of environmental value, σ_a^2 is additive genetic variance, $\mathbf{I}$ is identity matrix, and $\mathbf{A}$ is additive genetic relationship matrix.

The prior distributions of the variances of σ_u^2 , σ_a^2 and σ_e^2 are assumed to be scaled inverse χ^2 distributions:

$$p(\sigma_i^2) \propto (\sigma_i^2)^{-(\frac{\nu_i}{2}+1)} \exp\left[-\frac{\nu_i s_i^2}{2\sigma_i^2}\right], \quad i = u, a, e,$$

where ν_i is degree of freedom and s_i is a scale parameter.

The joint posterior distribution of all the parameters is

$$\begin{aligned} p(\mathbf{b}, \boldsymbol{\gamma}, \mathbf{u}, \mathbf{a}, \sigma_u^2, \sigma_a^2, \sigma_e^2 | \mathbf{y}) \\ \propto p(\mathbf{y} | \mathbf{b}, \boldsymbol{\gamma}, \mathbf{u}, \mathbf{a}, \sigma_e^2) p(\sigma_u^2) p(\mathbf{u} | \sigma_u^2) p(\sigma_a^2) \\ p(\mathbf{a} | \sigma_a^2) p(\sigma_e^2). \end{aligned} \quad [5]$$

Fully Conditional Posterior Distribution of the Location Parameters $\boldsymbol{\theta}$

Let $\boldsymbol{\theta}$ be the vector of all location parameters except for $\mathbf{u}$ [i.e., $\boldsymbol{\theta} = (\mathbf{b}', \boldsymbol{\gamma}', \mathbf{a}')'$], the fully conditional posterior distribution of $\boldsymbol{\theta}$ is:

$$p(\boldsymbol{\theta} | \mathbf{y}, \mathbf{u}, \sigma_u^2, \sigma_a^2, \sigma_e^2) \propto p(\mathbf{y} | \boldsymbol{\theta}, \mathbf{u}, \sigma_e^2) p(\mathbf{a} | \sigma_a^2). \quad [6]$$

Further, assuming $\mathbf{u}$ known, define:

$$\mathbf{y}_\theta = \mathbf{y} - \mathbf{Z}_u \mathbf{u} = \mathbf{X}\mathbf{b} + \mathbf{H}\boldsymbol{\gamma} + \mathbf{Z}_a \mathbf{a} + \mathbf{e}. \quad [7]$$

Thus, the fully conditional posterior distribution of $\boldsymbol{\theta}$ can be expressed as:

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{y}, \mathbf{u}, \sigma_u^2, \sigma_a^2, \sigma_e^2) &= p(\boldsymbol{\theta} | \mathbf{y}_\theta, \sigma_a^2, \sigma_e^2) \\ &\propto p(\mathbf{y}_\theta | \boldsymbol{\theta}, \sigma_e^2) p(\mathbf{a} | \sigma_a^2). \end{aligned} \quad [8]$$

Write the mixed model equations associated with [7] as $\mathbf{C}\hat{\boldsymbol{\theta}} = \mathbf{r}_\theta$. It can be shown that posterior distribution of location parameters $\boldsymbol{\theta}$, given dispersion parameters and $\mathbf{u}$, is multivariate normal:

$$\boldsymbol{\theta} \mid \mathbf{y}, \mathbf{u}, \sigma_u^2, \sigma_a^2, \sigma_e^2 \sim N(\hat{\boldsymbol{\theta}}, \mathbf{C}_0^{-1}). \quad [9]$$

Letting $\boldsymbol{\theta}_i$ denote an arbitrary element (or set of elements) of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_{-i}$ denote the vector of $\boldsymbol{\theta}$ excluding $\boldsymbol{\theta}_i$, the fully conditional posterior distribution of $\boldsymbol{\theta}_i$ is:

$$\boldsymbol{\theta}_i \mid \mathbf{y}, \mathbf{u}, \boldsymbol{\theta}_{-i}, \sigma_u^2, \sigma_a^2, \sigma_e^2 \sim N\left\{\mathbf{C}_{\boldsymbol{\theta}(i,i)}^{-1}\left[\mathbf{r}_{\boldsymbol{\theta}_i} - \mathbf{C}_{\boldsymbol{\theta}(i,-i)}\boldsymbol{\theta}_{-i}\right], \mathbf{C}_{\boldsymbol{\theta}(i,i)}^{-1}\right\}. \quad [10]$$

Fully Conditional Posterior Distribution of $\mathbf{u}$

The density of the fully conditional posterior distribution of $\mathbf{u}$ is:

$$p(\mathbf{u} \mid \mathbf{y}, \boldsymbol{\theta}, \sigma_u^2, \sigma_a^2, \sigma_e^2) \propto p(\mathbf{y} \mid \boldsymbol{\theta}, \mathbf{u}, \sigma_e^2)p(\mathbf{u} \mid \sigma_u^2). \quad [11]$$

Reparameterization for Linear $H \times E$ Model.

For a linear $H \times E$ model, an observation y can be described as:

$$\begin{aligned} y_{ij} &= \mathbf{x}'\mathbf{b} + u_i + \gamma_1 h_j u_i + a_i + e_{ij} \\ &= \mathbf{x}'\mathbf{b} + (1 + \gamma_1 h_j)u_i + a_i + e_{ij}. \end{aligned}$$

An alternative formulation of the reaction norm model [4] is:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_u^* \mathbf{u} + \mathbf{Z}_a \mathbf{a} + \mathbf{e}, \quad [12]$$

where $\mathbf{Z}_u^*$ is the incidence matrix obtained by replacing the nonzero element in the j th row of matrix $\mathbf{Z}_u$ with $(1 + \gamma_1 h_j)$.

Reparameterization for Nonlinear $H \times E$ Model. In the case of a nonlinear $H \times E$ model, the reparameterization can be done as follows:

$$\begin{aligned} y_{ij} &= \mathbf{x}'\mathbf{b} + u_i + \gamma_k h_j \phi_k(u_i) + \gamma_{k+1} h_j \phi_{k+1}(u_i) + a_j + e_{ij} \\ &= \mathbf{x}'\mathbf{b} + u_i + \gamma_k h_j \left(\frac{T_{k+1} - u_i}{T_{k+1} - T_k} \right) + \\ &\quad \gamma_{k+1} h_j \left(1 - \frac{T_{k+1} - u_i}{T_{k+1} - T_k} \right) + a_j + e_{ij} \\ &= \mathbf{x}'\mathbf{b} + \left(\frac{\gamma_k h_j T_{k+1} - \gamma_{k+1} h_j T_k}{T_{k+1} - T_k} \right) + \\ &\quad \left(1 + \frac{\gamma_{k+1} h_j - \gamma_k h_j}{T_{k+1} - T_k} \right) u_i + a_j + e_{ij}. \end{aligned}$$

Thus, the reaction norm model [4] for a nonlinear $H \times E$ interaction has an alternative form as:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{t} + \mathbf{Z}_u^* \mathbf{u} + \mathbf{Z}_a \mathbf{a} + \mathbf{e}, \quad [13]$$

where $\mathbf{t}$ is the vector with element

$$\left(\frac{\gamma_k h_j T_{k+1} - \gamma_{k+1} h_j T_k}{T_{k+1} - T_k} \right),$$

$\mathbf{Z}_u^*$ is the incidence matrix obtained by replacing the nonzero element in the j th row of matrix $\mathbf{Z}_u$ with

$$\left(1 + \frac{\gamma_{k+1} h_j - \gamma_k h_j}{T_{k+1} - T_k} \right).$$

Without knowing the range of environmental values, it is difficult to assign a set of reasonable values for the knots using the original scale. To simplify the implementation, the current model uses predetermined knot values in units of standard deviation of environment effects, say $\mathbf{T}^*$, the knot values in the original scale are assigned dynamically during Gibbs sampling process as:

$$\mathbf{T}_{(i)} = \mathbf{T}^* \sigma_{u(i)},$$

where $\sigma_{u(i)}$ is the standard deviation of $\mathbf{u}$ in the i th iteration.

Fully Conditional Distribution. Assuming that $\boldsymbol{\theta}$ is known, for a linear reaction norm model, define:

$$\mathbf{y}_u = \mathbf{y} - \mathbf{X}\mathbf{b} - \mathbf{Z}_a \mathbf{a} = \mathbf{Z}_u^* \mathbf{u} + \mathbf{e}. \quad [14]$$

For a nonlinear reaction norm model, define:

$$\mathbf{y}_u = \mathbf{y} - \mathbf{X}\mathbf{b} - \mathbf{t} - \mathbf{Z}_a \mathbf{a} = \mathbf{Z}_u^* \mathbf{u} + \mathbf{e}. \quad [15]$$

Because $p(\mathbf{y} \mid \boldsymbol{\theta}, \mathbf{u}, \sigma_u^2, \sigma_a^2, \sigma_e^2) = p(\mathbf{y}_u \mid \mathbf{u}, \sigma_u^2, \sigma_e^2)$, the fully conditional posterior distribution of $\mathbf{u}$ is:

$$\begin{aligned} p(\mathbf{u} \mid \mathbf{y}, \boldsymbol{\theta}, \sigma_u^2, \sigma_a^2, \sigma_e^2) &= p(\mathbf{u} \mid \mathbf{y}_u, \sigma_u^2, \sigma_e^2) \\ &\propto p(\mathbf{y}_u \mid \mathbf{u}, \sigma_e^2)p(\mathbf{u} \mid \sigma_u^2). \end{aligned} \quad [16]$$

Write the mixed model equations associated with [14] and [15] as $\mathbf{C}_u \hat{\mathbf{u}} = \mathbf{r}_u$. Then the fully conditional distribution of $\mathbf{u}$ is:

$$\mathbf{u} \mid \mathbf{y}, \boldsymbol{\theta}, \sigma_u^2, \sigma_a^2, \sigma_e^2 \sim N(\hat{\mathbf{u}}, \mathbf{C}_u^{-1}), \quad [17]$$

and for the i th element of $\mathbf{u}$, the fully conditional posterior distribution is:

$$\mathbf{u}_i \mid \mathbf{y}, \boldsymbol{\theta}, \mathbf{u}_{-i}, \sigma_u^2, \sigma_a^2, \sigma_e^2 \sim N\left\{\mathbf{C}_{u(i,i)}^{-1} \left[\mathbf{r}_{u_i} - \mathbf{C}_{u(i,-i)} \mathbf{u}_{-i}\right], \mathbf{C}_{u(i,i)}^{-1}\right\}. \quad [18]$$

Fully Conditional Posterior Distribution of Dispersion Parameters

The fully conditional posterior distribution of disperse parameters are recognized as scaled inverse χ^2 distributions. Thus,

$$\sigma_u^2 \sim \text{Inv} - X^2 \left[\nu_u + n_u, (\mathbf{u}'\mathbf{u} + \nu_u s_u^2) / (\nu_u + n_u) \right]$$

$$\sigma_a^2 \sim \text{Inv} - X^2 \left[\nu_a + n_a, (\mathbf{a}'\mathbf{A}^{-1}\mathbf{a} + \nu_a s_a^2) / (\nu_a + n_a) \right]$$

$$\sigma_e^2 \sim \text{Inv} - X^2 \left\{ \nu_e + n, \left(\begin{array}{c} (\mathbf{y} - \mathbf{Xb} - \mathbf{H}\boldsymbol{\gamma} \\ - \mathbf{Z}_u \mathbf{u} - \mathbf{Z}_a \mathbf{a})' \\ (\mathbf{y} - \mathbf{Xb} - \mathbf{H}\boldsymbol{\gamma} \\ - \mathbf{Z}_u \mathbf{u} - \mathbf{Z}_a \mathbf{a}) \\ + \nu_e s_e^2 \end{array} \right) / (\nu_e + n) \right\},$$

where ν_i ($i = e, a, u$) is the degree of freedom and s_i^2 is the scale parameter for the prior distribution of σ_i^2 , n is the number of observations, n_u is the order of $\mathbf{u}$, and n_a is the order of $\mathbf{a}$.

Implementation of the Gibbs Sampler

The algorithm of Gibbs sampler for the proposed reaction norm model is as follows:

1. Provide initial σ_u^2, σ_a^2 , and σ_e^2 . For a nonlinear H $\times$ E model, provide also knots (can be expressed as standard unit of $\mathbf{u}$).
2. Compute $\mathbf{y}_u, \mathbf{Z}_u^*, \mathbf{C}_u$, and $\mathbf{r}_u$ and draw $\mathbf{u}_i$ from $N\left\{\mathbf{C}_{u(i,i)}^{-1} \left[\mathbf{r}_{u_i} - \mathbf{C}_{u(i,-i)} \mathbf{u}_{-i}\right], \mathbf{C}_{u(i,i)}^{-1}\right\}$.
3. Compute $\mathbf{y}_\theta, \mathbf{C}_\theta$, and $\mathbf{r}_\theta$ and draw $\boldsymbol{\theta}_i$ from $N\left\{\mathbf{C}_{\theta(i,i)}^{-1} \left[\mathbf{r}_{\theta_i} - \mathbf{C}_{\theta(i,-i)} \boldsymbol{\theta}_{-i}\right], \mathbf{C}_{\theta(i,i)}^{-1}\right\}$.
4. Sample new σ_u^2, σ_a^2 , and σ_e^2 from scaled inverse χ^2 distributions.
5. Repeat steps 2 to 4 until enough samples are obtained.

The computing program for the proposed model was developed and integrated into the DMU-package (Madsen and Jensen, 2004; Madsen et al., 2006). In the actual implementation, the “iteration on data” technique, which avoids storing $\mathbf{C}_\theta$ and $\mathbf{C}_u$, was applied.

Simulation Studies

Data Generation. The proposed method was evaluated using a simulation study. Data were generated according to a structure mimicking dairy cattle but simplifying the method of selection and replacement.

The data were produced in 2 steps. In the first step, data were generated without considering herd-year effect and heterosis. An observation of base generation was generated using a model:

$$y_0 = \mu + a + e,$$

where $a \sim N(0, \sigma_a^2)$, and $e \sim N(0, \sigma_e^2)$.

In the following generations, an observation was simulated as:

$$y_0 = \mu + 0.5(a_{sire} + a_{dam}) + \varepsilon + e,$$

where $\varepsilon \sim N(0, 0.5\sigma_a^2)$.

Two populations were created for a period of 20 yr. The cows and young bulls were selected randomly. The proven bulls were selected on the phenotypic means of their daughters. From yr 3 and onward, approximately 25% of the cows in population I were inseminated with the semen imported from population II, but all daughters of young bulls were used for replacement. Cows did not move from one herd to another. The breeding scheme for the 2 populations is shown in Table 1.

At the second step, herd-year effect was generated from $u \sim N(0, \sigma_u^2)$ and total heterosis was generated as a function of breed heterozygosity (*het*) and herd-year effect (u) as:

$$\begin{aligned} \text{Heterosis} &= 20 \times \text{het} + 4 \times \text{het} \times u / \sigma_u \\ &- 2.5 \times \text{het} \times (u / \sigma_u)^2 - 0.25 \times \text{het} \times (u / \sigma_u)^3. \end{aligned}$$

Thus, a final observation was generated as:

$$y = y_0 + u + \text{heterosis}.$$

The simulation study included 4 scenarios with regard to the size of additive genetic and herd-year variation (M = moderate variation and S = small variation in the context). For each scenario, 10 replicates were generated. The parameters used in the 4 scenarios are

Table 1. Breeding scheme for the 2 simulated populations

Item	Population I	Population II
No. of bulls tested per year	50	50
No. of daughters per test bull	50	100
No. of proven bull use per year	50	50
Replacement rate of cow (%)	30	30
Replacement rate of sire (%)	30	30
No. of herds per year	100	100
No. of cows per herd-year	150	300
No. of first-parity records per herd-year	50	100

reported in Table 2. Data for the first parity of population I were used in the current analysis.

To simplify simulation procedure, Mendelian sampling term was generated from a distribution $\varepsilon \sim N(0, 0.5\sigma_a^2)$, instead of true distribution $\varepsilon \sim N[0, 0.5(1 - F_p)\sigma_a^2]$, where $F_p = 0.5(F_s + F_d)$ is the average of parent inbreeding coefficients. To evaluate the feasibility of ignoring inbreeding, a data set was simulated according to the present population structure with $\sigma_a^2 = 120$ and $\sigma_e^2 = 280$. Based on this data set, F_p ranged from 0 to 0.158 with mean of 0.001. For the individual in the last birth year, the mean of F_p was 0.003. The proportion of F_p over 0.02 was 1% in the whole population and 4% in the last birth year. The proportion of F_p over 0.05 was 0.25% in the whole population and 0.55% in the last year. These suggested that the effect of ignoring inbreeding on simulated data could be negligible.

Statistical Analysis

The simulated data were analyzed using 4 alternative models. Model 1 ignored $H \times E$ interaction,

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_{bp}\gamma_{bp} + \mathbf{H}^0\gamma_0 + \mathbf{Z}_u\mathbf{u} + \mathbf{Z}_a\mathbf{a} + \mathbf{e},$$

where γ_{bp} was the regression on breed proportion (population II), γ_0 was the regression coefficient on breed heterozygosity ($\mathbf{H}^0$), $\mathbf{u}$ was the vector of random herd-year effect, $\mathbf{a}$ was the vector of additive genetic effects, and $\mathbf{e}$ was the vector of residuals.

Model 2 used the true curve function and true covariates ($\mathbf{H}^t$) to fit heterosis, and treating herd-year ($\mathbf{u}$) as random effect,

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_{bp}\gamma_{bp} + \mathbf{H}^t\boldsymbol{\gamma} + \mathbf{Z}_u\mathbf{u} + \mathbf{Z}_a\mathbf{a} + \mathbf{e},$$

where $\boldsymbol{\gamma}$ was the vector of regression coefficients on known covariates of the reaction norm [i.e., the first, second, third, and fourth column of $\mathbf{H}^t$ was breed heterozygosity (het), $het \times u_t / \sigma_{u_t}$, $het \times (u_t / \sigma_{u_t})^2$, and

$het \times (u_t / \sigma_{u_t})^3$, respectively, where u_t is true herd-year effect].

Model 3 was the same as model 2 but treating herd-year effect as fixed effect,

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_{bp}\gamma_{bp} + \mathbf{H}^t\boldsymbol{\gamma} + \mathbf{X}_u\mathbf{u} + \mathbf{Z}_a\mathbf{a} + \mathbf{e}.$$

Model 4 (the proposed model) was the model with unknown covariates ($\mathbf{H}^u$) of the reaction norm for heterosis, and treating herd-year ($\mathbf{u}$) as random effect,

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}_{bp}\gamma_{bp} + \mathbf{H}^u\boldsymbol{\gamma} + \mathbf{Z}_u\mathbf{u} + \mathbf{Z}_a\mathbf{a} + \mathbf{e},$$

where $\boldsymbol{\gamma}$ was the vector of spline regression coefficients on unknown covariates of the reaction norm [i.e., regressions on $het \times \phi_k(\hat{u})$]. Eight knots were used in the spline regression with the values in unit of standard deviation as $(-2.0, -1.2, -0.7, -0.3, 0, 0.3, 0.7, 1.2, 2.0)$.

RESULTS

Based on the parameters in the simulation, the total heterosis was different in various herd-years (Figure 1). At a breed heterozygosity of 100%, the total heterosis was 4 in the herd-year with an effect of -2 units of standard deviation, 20 in the herd-year with an effect equal to the mean of herd-year effect, and 16 in the herd-year with an effect of 2 units of standard deviation. The maximum heterosis was 21.5 in the herd-year with an effect of 0.72 units of standard deviation.

Some realized statistics of the simulated data are reported in Table 3. Averaged over 10 replicates, the mean breed heterozygosity was 0.5 in the last year and 0.356 across all years. Defining phenotypic variance as the sum of additive genetic variance and residual variance, the proportion of total heterosis variance to phenotypic variance was 9.6%, and the proportion of

Table 2. Additive genetic variance (σ_a^2), variance of herd-year effects (σ_u^2), and residual variance (σ_e^2) used in the 4 scenarios of data simulation

Scenario ¹	σ_a^2	σ_u^2	σ_e^2
MM	120	80	280
MS	120	20	280
SM	20	80	380
SS	20	20	380

¹MM = moderate variation of additive genetic effect and herd-year effect; MS = moderate variation of additive genetic effect and small variation of herd-year effect; SM = small variation of additive genetic effect and moderate variation of herd-year effect; SS = small variation of additive genetic effect and herd-year effect.

Table 3. Realized breed heterozygosity (*het*) in the last year, mean *het*, variance of total heterosis (σ^2_{HT}), variance of environment-dependent heterosis ($\sigma^2_{H \times E}$), within-herd-year variance of heterosis $\times$ environment ($H \times E$) [$\sigma^2_{W(H \times E)}$], sire variance of $H \times E$ corrected for herd-year mean of $H \times E$ [$\sigma^2_{S(H \times E)}$], correlation between additive genetic effect and $H \times E$ [$r_{(a, H \times E)}$], and correlation between daughter's $H \times E$ and dam's $H \times E$ [$r_{(H \times Eo, H \times Ed)}$], averaged over 10 replicates

Scenario ¹	Last year <i>het</i>	Mean <i>het</i>	σ^2_{HT}	$\sigma^2_{H \times E}$	$\sigma^2_{W(H \times E)}$	$\sigma^2_{S(H \times E)}$	$r_{(a, H \times E)}$	$r_{(H \times Eo, H \times Ed)}$
MM	0.498	0.356	38.40	6.292	2.534	0.397	−0.048	0.028
MS	0.500	0.356	38.47	6.300	2.538	0.398	−0.034	0.028
SM	0.498	0.356	38.40	6.292	2.544	0.397	−0.048	0.028
SS	0.500	0.356	38.47	6.300	2.550	0.398	−0.034	0.028

¹MM = moderate variation of additive genetic effect and herd-year effect; MS = moderate variation of additive genetic effect and small variation of herd-year effect; SM = small variation of additive genetic effect and moderate variation of herd-year effect; SS = small variation of additive genetic effect and herd-year effect.

environment-dependent heterosis variance ($H \times E$ variance) was 1.58%. The within-herd-year $H \times E$ variance was only about 0.64% of phenotypic variance. Further, after adjusting for the herd-year mean of environment-dependent heterosis, the variance between sires for the adjusted environment-dependent heterosis was very small (0.4). In addition, there was no correlation between additive genetic effect and environment-dependent heterosis, or between dam and daughter's environment-dependent heterosis in the simulated data.

The correlation between true heterosis and predicted heterosis using the proposed model (model 4) was 0.990 for the scenario with moderate variation of herd-year effects and 0.982 for the scenario with small herd-year variation (Table 4). The accuracies were even higher than those obtained by using true environmental covariable and treating herd-year as random effect (model 2) but slightly lower than those using true environmental covariable and treating herd-year as a fixed effect (model 3). As expected, the model ignoring $H \times E$

interaction (model 1) had the poorest prediction of heterosis.

The correlations between true breeding value and predicted breeding values were almost the same among different models (Table 5). Ignoring $H \times E$ interaction did not significantly reduce the accuracy of predicted breeding value based on the simulated data. On the other hand, the accuracy achieved from the model using true environmental covariable and treating herd-year as fixed effect was slightly lower than the other models. In line with the amount of additive genetic variation, the accuracy of predicted breeding value was higher in scenarios with moderate additive genetic variation than scenarios with small additive genetic variation.

DISCUSSION

The presented study proposed a reaction norm model to analyze $H \times E$ interaction with unknown environmental covariate. The results from a simulation study showed that the model and method led to adequate inferences of heterosis and breeding value. On the other hand, ignoring $H \times E$ interaction led to a poor prediction of heterosis but did not harm the prediction of breeding value, based on the simulated data.

In the proposed reaction norm model, total heterosis consists of an environment-independent heterosis and an environment-dependent heterosis. The environment-dependent heterosis was described as a function of breed heterozygosity and environmental value (e.g., herd-year effect). Reaction norm model has the advantage of dealing with an infinite number of environment levels with few parameters. It is useful when phenotypes change gradually and continuously over an environmental gradient (de Jong, 1995). Although reaction norm models have been widely used to analyze additive genetic effect $\times$ environment interaction in dairy cattle (Calus et al., 2002; Kolmodin et al., 2002; Shariati et al., 2007), the present study is the first to propose a reaction norm model to analyze $H \times E$ interaction.

Figure 1. Heterosis in relation to environmental value (in units of SD) in the simulation, for breed heterozygosity (*het*) at 0.5 and 1.0.

Table 4. Correlation between true total heterosis and predicted total heterosis using various models,¹ averaged over 10 replicates

Scenarios ²	Model 1	Model 2	Model 3	Model 4
MM	0.924	0.972	0.999	0.990
MS	0.924	0.946	0.999	0.983
SM	0.924	0.967	0.999	0.990
SS	0.924	0.944	0.999	0.982

¹Model 1 = ignoring heterosis $\times$ environment interaction; model 2 = with true environmental covariable, treating herd-year as random effect; model 3 = with true environmental covariable, treating herd-year as fixed effect; model 4 = with unknown environmental covariable and spline regression (the proposed approach).

²MM = moderate variation of additive genetic effect and herd-year effect; MS = moderate variation of additive genetic effect and small variation of herd-year effect; SM = small variation of additive genetic effect and moderate variation of herd-year effect; SS = small variation of additive genetic effect and herd-year effect.

Various known descriptive variables (e.g., temperature, humidity, feeding level) can be used as covariates of reaction norm, but a common used environmental gradient in practical genetic evaluation is the effect of production environment. However, the effect of production environment is unknown. One approximation is to use the average phenotypic performance of the individuals in a given environment as environmental covariate (Kolmodin et al., 2002). Another approximation is to estimate effect of production environment using a simple linear model or a standard additive genetic model, and then use the estimate as proxy of environmental covariate (Calus et al., 2002; Oseni et al., 2004). These approximations could be unsatisfactory in many situations, especially when the trait of interest is under selection (Calus et al., 2004; Su et al., 2006). Su et al. (2006) developed a Bayesian reaction norm model that allows inferring environmental covariates and the other parameters in the model simultaneously. The advantage of this approach over the approximate methods has been shown in a simulation study (Su et al., 2006). Therefore the proposed model applies the methodology reported by Su et al. (2006).

The proposed model applies linear spline regression to fit nonlinear reaction form. Spline regression has the advantage of accommodating various nonlinear reac-

tion norms including nonparametric reaction norms. One challenge in spline regression is to choose the right number and locations of the knots. Too many knots will increase complexity and the uncertainty of the estimates, whereas too few would reduce the predictive ability. Usually, the number and positions of knots in splines are chosen corresponding to the pattern of the trajectory. Knots can be dense in regions with fast changes and sparse in regions with slow changes. The strategies of choosing number and positions of knots have been reported by several studies (Biller, 2000; Zhou and Shen, 2001; Hansen and Kooperburg, 2002; Misztal, 2006). With unknown environmental covariates, it is difficult to predetermine the positions of knots. The proposed model uses predetermined values of knots in standard deviation units of environmental values. The positions of knots in original scale are assigned dynamically during the process of Gibbs sampling. The results from the simulation study show that this technique works well.

When applying the proposed model to analyze $H \times E$ interaction, attention should be paid to interpretation of the regression coefficients in relation to the scale of the covariables. In the proposed model, heterosis is described as $heterosis = \left[\beta_0 + \sum_{k=1}^m \beta_k \phi_k(u_i) \right] h$. According

Table 5. Correlation between true breeding value and predicted breeding value using various models,¹ averaged over 10 replicates

Scenarios ²	Model 1	Model 2	Model 3	Model 4
MM	0.714	0.715	0.709	0.715
MS	0.715	0.717	0.709	0.716
SM	0.501	0.504	0.494	0.509
SS	0.506	0.510	0.494	0.509

¹Model 1 = ignoring heterosis $\times$ environment interaction; model 2 = with true environmental covariable, treating herd-year as random effect; model 3 = with true environmental covariable, treating herd-year as fixed effect; model 4 = with unknown environmental covariable and spline regression (the proposed approach).

²MM = moderate variation of additive genetic effect and herd-year effect; MS = moderate variation of additive genetic effect and small variation of herd-year effect; SM = small variation of additive genetic effect and moderate variation of herd-year effect; SS = small variation of additive genetic effect and herd-year effect.

to the properties of regression coefficient, in general, scaling environmental value (u) by a constant (e.g., by SD) will change β_k but not β_0 , and scaling heterozygosity (h) by a constant (e.g., h in percentage) will change both β_k and β_0 . Moving the origin of u (e.g., subtracting the mean) will change β_0 or both β_k and β_0 , dependent on the feature of $\phi_k(u_i)$. However, these changes in scale of u and h will not change the prediction of heterosis. On the other hand, moving the origin of h (e.g., subtracting a constant c) will result in biased prediction of heterosis with an amount of $\beta_0 c$ for all individuals and $\sum_{k=1}^m \beta_k \phi_k(u_i) c$, which is different for individuals in different environments.

The simulation study showed that the proposed model detected $H \times E$ interaction and predicted heterosis and breeding value accurately. The accuracy of predicted heterosis was even higher than the model using true environmental values and true curve function, and treating herd-year as a random effect (Table 4). In addition, the accuracy of predicted breeding values from the model treating herd-year as fixed effect was lower than the models treating herd-year as random effect (Table 5). It has been shown that treating herd-year effect as random can increase the accuracy of selection (Visscher and Goddard, 1993) and the predictive ability of the model (Babot et al., 2003). On the other hand, in the situation of an association between sire effect and herd effect or an environmental trend, treating herd-year effect as random could lead to biased estimates of breeding values (Henderson, 1973; Visscher and Goddard, 1993; Babot et al., 2003). However, Ugarte et al. (1992) reported that even under an association between sire and herd, there is no bias in predicted breeding values in the simulated scenarios where sizes of contemporary groups are larger than 12.

Based on the simulated data, a model ignoring $H \times E$ interaction does not harm the accuracy of predicted breeding value. This can be explained from 2 aspects. First, although there was a considerable $H \times E$ interaction in the simulated populations, the variance of environment-dependent component of heterosis was small. In addition, though there was a large difference in the levels of breed heterozygosity between years, the variation within herd-year was small because sires were used randomly. In a model ignoring $H \times E$ interaction, the between-herd-year variation of $H \times E$ interaction (the major part of $H \times E$ variation) was accounted by herd-year effect. It indicates that the random error due to $H \times E$ interaction is negligible in the current simulated data.

The second reason is that in the simulated scenarios (sires were used randomly across environments), ignor-

ing $H \times E$ interaction does not lead to systematic bias in prediction of breeding value. As shown in Table 3, the correlation between $H \times E$ effect and breeding value, and between dam and daughter's $H \times E$ effect, is close to zero. The variance between sires for $H \times E$ effect, after correction for herd-year mean of $H \times E$ interaction, is extremely small. These results indicate that ignoring $H \times E$ interaction does not increase or decrease the resemblance between relatives. Therefore, the systematic error due to ignoring $H \times E$ is negligible.

However, ignoring $H \times E$ in genetic evaluation might lead to a serious problem in situations other than those mimicked by the current simulation. In the case with many close relatives with similar heterozygosities within the same herd-year, ignoring $H \times E$ could lead to a systematic error in prediction of breeding values. It would be useful to carry out further studies on the influence of ignoring $H \times E$ on prediction of breeding value, with regard to various breeding schemes and population structures.

The proposed model and method were developed to handle $H \times E$ interaction using a reaction norm with few parameters. This approach could be a good alternative to estimate $H \times E$ interaction and predict breeding value in situations where there are a large number of environment levels and the environments can be qualified as an environmental gradient, but not necessary to perform well in all situations. For example, to analyze environment-dependent heterosis in few levels of heterozygosity and few levels of environment, the traditional classified interaction model could be the best choice.

The weakness of this study was that there were some flaws in simulating data. In the simulation study, data were generated to mimic dairy cattle populations. However, the properties of the simulation were not really agreeable with real-life scenarios, due to simplifying the simulation procedure. First, the simulation did not generate history generations to create the difference in allele frequencies between 2 populations, and a certain level of inbreeding. Second, the Mendelian sampling terms for offspring were generated without considering inbreeding coefficients of parents. Third, the function to create heterosis leads to a negative heterosis in environments with a value less than -2.4 standard deviation, which could be not realistic.

CONCLUSIONS

The present study is the first to propose a reaction norm model to analyze $H \times E$ interaction. The proposed model applies a Bayesian approach, which allows estimating environmental covariate and the other

parameters of the model simultaneously. A simulation study shows that the proposed model detects $H \times E$ interaction and predicts heterosis and breeding value accurately. Therefore, the proposed model could be a good approach to estimate $H \times E$ interaction and predict breeding value in situations in which heterosis changes gradually and continuously over an environmental gradient.

ACKNOWLEDGMENTS

We thank the Danish Cattle Federation (Aarhus, Denmark) for funding this research.

REFERENCES

- Arthur, P. F., H. Hearnshaw, and P. D. Stephenson. 1999. Direct and maternal additive and heterosis effects from crossing *Bos indicus* and *Bos taurus* cattle: Cow and calf performance in two environments. *Livest. Prod. Sci.* 57:231–241.
- Babot, D., J. L. Noguera, L. Alfonso, and J. Estany. 2003. Fixed or random contemporary groups in genetic evaluation for litter size in pigs using a single trait repeatability animal model. *J. Anim. Breed. Genet.* 120:12–22.
- Barlow, R. 1981. Experimental evidence for interaction between heterosis and environment in animals. *Anim. Breed. Abstr.* 49:715–737.
- Biller, C. 2000. Adaptive Bayesian regression splines in semiparametric generalized linear models. *J. Comput. Graph. Statist.* 9:122–140.
- Brown, M. A., A. H. Brown Jr., W. G. Jackson, and J. R. Miesner. 2001. Genotype $\times$ environment interactions in milk yield and quality in Angus, Brahman, and reciprocal-cross cows on different forage systems. *J. Anim. Sci.* 79:1643–1649.
- Bryant, J. R., N. López-Villalobos, J. E. Pryce, C. W. Holmes, D. L. Johnson, and D. J. Garrick. 2007. Environmental sensitivity in New Zealand dairy cattle. *J. Dairy Sci.* 90:1538–1547.
- Calus, M. P. L., P. Bijma, and R. F. Veerkamp. 2004. Effects of data structure on the estimation of covariance functions to describe genotype by environment interaction in a reaction norm model. *Genet. Sel. Evol.* 36:489–507.
- Calus, M. P. L., A. F. Groen, and G. de Jong. 2002. Genotype $\times$ environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *J. Dairy Sci.* 85:3115–3123.
- de Jong, G. 1995. Phenotypic plasticity as a product of selection in a variable environment. *Am. Nat.* 145:493–512.
- Hansen, M., and C. Kooperburg. 2002. Spline adaptation in extended linear models. *Stat. Sci.* 17:1–51.
- Henderson, C. R. 1973. Sire evaluation and genetic trends. Pages 10–41 in *Proc. of the Animal Breeding and Genetics Symposium in Honor of Dr. J. L. Lush*. Am. Soc. Anim. Sci., Champaign IL.
- Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agric. Scand., Sect. Anim. Sci.* 52:11–24.
- Madsen, P., and J. Jensen. 2004. A User's Guide to DMU, Version 6, Release 4.5. Danish Institute of Agricultural Sciences, Tjele, Denmark.
- Madsen, P., P. Sørensen, G. Su, L. H. Damgaard, H. Thomsen, and R. Labouriau. 2006. DMU – A package for analyzing multivariate mixed models. *Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Brazil*. CD-ROM Communication No. 07-06.
- Misztal, I. 2006. Properties of random regression models using linear splines. *J. Anim. Breed. Genet.* 123:74–80.
- Oseni, S., I. Misztal, S. Tsuruta, and R. Rekaya. 2004. Genetic components of days open under heat stress. *J. Dairy Sci.* 87:3022–3028.
- Phillips, W. A., M. A. Brown, A. H. Brown Jr., and S. W. Coleman. 2001. Genotype $\times$ environment interactions for postweaning performance in crossbred calves grazing winter wheat pasture or dormant native prairie. *J. Anim. Sci.* 79:1370–1377.
- Shariati, M. M., G. Su, P. Madsen, and D. Sørensen. 2007. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. *J. Dairy Sci.* 90:5759–5766.
- Strandberg, E. 2006. Analysis of genotype by environment interaction using random regression models. *Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Brazil*. CD-ROM Communication No. 25-05.
- Su, G., P. Madsen, M. S. Lund, D. Sørensen, I. R. Korsgaard, and J. Jensen. 2006. Bayesian analysis of the linear reaction norm model with unknown covariate. *J. Anim. Sci.* 84:1651–1657.
- Ugarte, E., R. Alenda, and M. J. Carabano. 1992. Fixed or random contemporary groups in genetic evaluations. *J. Dairy Sci.* 75:269–278.
- Visscher, P. M., and M. E. Goddard. 1993. Fixed and random contemporary groups. *J. Dairy Sci.* 76:1444–1454.
- Zhou, S., and X. Shen. 2001. Spatially adaptive regression splines and accurate knot selection schemes. *J. Am. Stat. Assoc.* 96:247–259.