Published August 30, 2016

RESEARCH

A Hierarchical Bayesian Estimation Model for Multienvironment Plant Breeding Trials in Successive Years

Diego Jarquín, Sergio Pérez-Elizalde, Juan Burgueño, and José Crossa*

ABSTRACT

In agriculture and plant breeding, multienvironment trials over multiple years are conducted to evaluate and predict genotypic performance under different environmental conditions and to analyze, study, and interpret genotype × environment interaction ($G \times E$). In this study, we propose a hierarchical Bayesian formulation of a linear–bilinear model, where the conditional conjugate prior for the bilinear (multiplicative) $G \times E$ term is the matrix von Mises–Fisher (mVMF) distribution (with environments and sites defined as synonymous). A hierarchical normal structure is assumed for linear effects of sites, and priors for precision parameters are assumed to follow gamma distributions. Bivariate highest posterior density (HPD) regions for the posterior multiplicative components of the interaction are shown within the usual biplots. Simulated and real maize (*Zea mays* L.) breeding multisite data sets were analyzed. Results showed that the proposed model facilitates identifying groups of genotypes and sites that cause $G \times E$ across years and within years, since the hierarchical Bayesian structure allows using plant breeding data from different years by borrowing information among them. This model offers the researcher valuable information about $G \times E$ patterns not only for each 1-yr period of the breeding trials but also for the general process that originates the response across these periods.

D. Jarquín, Dep. of Agronomy and Horticulture, Univ. of Nebraska–Lincoln, 321 Keim Hall, Lincoln, NE, 68503-0915; S. Pérez-Elizalde, Colegio de Postgraduados, Km. 36.5 Carretera México–Texcoco, Montecillo, Estado de México, 56230, Mexico; J. Burgueño, and J. Crossa, Biometrics and Statistics Unit (BSU), International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600, Mexico D.F., Mexico. Received 6 Aug. 2015. Accepted 17 Nov. 2015.*Corresponding author (j.crossa@cgiar.org).

Abbreviations: AMMI, additive main effects and multiplicative interaction; $G \times E$, genotype × environment interaction; HPD, highest posterior density; MCMC, Markov Chain Monte Carlo; mVMF, matrix von Mises–Fisher; SREG, site regression.

Linear–bilinear models are useful for analyzing two-way tables with interaction; they are applied in different areas of research, such as medicine, agriculture, social sciences, and engineering, where estimating the main effects of rows and columns, as well as their interactions, is critically important. When experiments do not have a specific treatment structure suitable for generating contrasts between rows and columns, general linear–bilinear models are useful for analyzing main effects and their interactions (Cornelius and Seyedsadr, 1997). Experiments with unstructured row and column factors are typically used in agriculture and plant breeding, where very costly field trials must be established and large numbers of genotypes (hybrids, lines, synthetics, etc.) are evaluated in multienvironment, multiyear trials. In these experiments, the mean response of rows (genotypes), columns (environments or sites), and their interaction ($G \times E$) may be modeled by combining linear–bilinear terms (Cornelius et al., 2001; Crossa et al., 2004).

Published in Crop Sci. 56:2260–2276 (2016).
doi: 10.2135/cropsci2015.08.0475

One linear–bilinear two-way model that is used in plant breeding trials for assessing the average response ($\overline{y}_{ij}$) of g genotypes ($i = 1,2,\ldots,g$) evaluated in s sites ($j = 1,2,\ldots,s$) is

$$\overline{y}_{ij} = \mu + \beta_j + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk} + \overline{\varepsilon}_{ij} \qquad [1]$$

where μ and β_j are the grand mean and the effect of the jth site, respectively. For ordinary least squares estimation, the linear terms (μ and β_j) are fitted in a first step, and then the bilinear term ($\sum_{k=1}^{t} \lambda_k u_{ik} v_{jk}$), which gives the combined effects of the ith genotype plus the ijth interaction, is fitted through singular value decomposition of the average residual $g \times s$ matrix (Gabriel, 1978). From where, λ_k is the singular value subject to $\lambda_1 \geq \ldots \lambda_t \geq 0$; u_{ik} and v_{jk} are the elements of the left and right singular vectors, respectively, subject to the orthonormality constraints $\sum_i u_{ik}^2 = \sum_j v_{jk}^2 = 1$ and for $k \neq k'$, $\sum_i u_{ik} u_{ik'} = \sum_j v_{jk} v_{jk'} = 0$ to enable identifiability. Additionally, it may be assumed that the errors $\overline{\varepsilon}_{ij} = \frac{\sum_{l=1}^{n_{ij}} \varepsilon_{ijl}}{n_{ij}}$ are identically and indepen dently normally distributed with mean 0 and variance $\sigma_\varepsilon^2 / n_{ij}$ (with $l = 1,2,\ldots,n_{ij}$, where n_{ij} is the number of replicates for each genotype–site–year combination); in what follows, an equal number of observations per cell ($n = n_{ij}$) is assumed. The number of multiplicative terms is $t \leq p$, where p is the number of terms required to saturate the model if fitted by least squares, namely, $p = \min(g,s) - 1$. It should be noted that the effect of the design (i.e., complete or incomplete blocks, row and column, etc.), as well as the covariates, could easily be incorporated into the model through the covariance structure on the linear terms.

In matrix notation, [1] can be expressed as

$$\mathbf{Y} = \mu \mathbf{1}_g \mathbf{1}_s' + \boldsymbol{\beta}' \otimes \mathbf{1}_g + \mathbf{U}\mathbf{D}\mathbf{V}' + \mathbf{E} \qquad [2]$$

where $\mathbf{Y} = \left[\overline{y}_{ij}\right]$ is a matrix of cell means of order $g \times s$, $\boldsymbol{\beta} = \left[\beta_j\right]$, $\mathbf{D} = \mathbf{diag}\left(\lambda_k, k = 1,2,\ldots,t\right)$, $\mathbf{U} = (\mathbf{u}_1,\ldots,\mathbf{u}_t)$, $\mathbf{u}_1 = [u_{ik}]$, $\mathbf{V} = (\mathbf{v}_1,\ldots,\mathbf{v}_t)$, $\mathbf{v}_i = [v_{ik}]$, and $\mathbf{E} = \left[\overline{\varepsilon}_{ij}\right]$. As previously mentioned, with ordinary least squares, the parameters in [2] are sequentially estimated by first fitting the linear terms and then estimating the bilinear terms as the first t components of the singular value decomposition of the residual matrix $\mathbf{Z} = \mathbf{Y} - \hat{\mu}\mathbf{1}_g\mathbf{1}_s' - \hat{\boldsymbol{\beta}}' \otimes \mathbf{1}_g$, where $\hat{\mu}$, and $\hat{\boldsymbol{\beta}}$ are the estimates obtained in the first step by ignoring the bilinear terms (Gabriel, 1978). Model [2], known as the site regression (SREG) model (Crossa and Cornelius, 1997), does not explicitly include the genotype main effect, which implies that this effect is nested within the site effects (β) and thus absorbed into the bilinear decomposition of the interaction term. Therefore, inference in model [2] refers to the main effects of genotypes plus the effects of $G \times E$. Another commonly used linear–bilinear model is the additive main effects and multiplicative interaction (AMMI) model that explicitly introduced the genotype main effect in [2]; thus the bilinear term of AMMI only fits the $G \times E$ term.

An initial seminal work for a singular value shrinkage estimator of a linear–bilinear AMMI model was proposed by Cornelius and Crossa (1999); it produced fitted models that were as good as, or better than, the best linear unbiased predictor obtained from random models. Recently, a Bayesian shrinkage estimator similar to the one proposed by Cornelius and Crossa (1999) was proposed by da Silva et al. (2015). However, until a few years ago, Bayes analysis of linear–bilinear models had been computationally limited by the complex structure of the parameter space because of the orthonormal bases used in singular value decomposition. Early Bayesian approaches to overcome this limitation were proposed by Viele and Srinivasan (2000) and Liu (2001). These approaches for sampling the conditional posterior distributions were performed within the vector framework, that is, for the joint posterior distribution of vectors $\mathbf{u}_k$ and $\mathbf{v}_k$, which are columns from $\mathbf{U}$ and $\mathbf{V}$, respectively. Crossa et al. (2011) used this vector approach to analyze real data from multienvironment plant breeding trials.

A generalization of the vector approach to a matrix approach for Bayesian inference was proposed by Perez-Elizalde et al. (2012) who used the mVMF distribution as the prior for the orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$. To obtain a Markov Chain Monte Carlo (MCMC) sample from conditional posterior distributions, which are mVMF distributions onto the multidimensional sphere, the authors used the algorithm proposed by Hoff (2009). The difficulties of working in an overparameterized framework and complying with AMMI model constraints were recently recognized by Josse et al. (2014); these authors presented an alternative approach to the Bayesian AMMI inference method proposed by Perez-Elizalde et al. (2012) that ignores the issue of overparameterization and model constraints at the prior level while applying appropriate processing at the posterior level.

Plant breeders often perform analyses of two-way tables (genotype × site or genotype × environment) in several consecutive periods (years) to find groups of genotypes and sites with $G \times E$ effects and detect stable genotypes across different environmental conditions and periods. Also, breeders usually have historical records of experimental data that can be incorporated into a hierarchical model involving several periods (years) of multisite or multienvironment trials. However, the scope of the current linear–bilinear models for analyzing plant

breeding trials that are established in multienvironments repeated across several years is limited. A common practice is to collapse environments (sites) and years into one unique factor; however, this causes a substantial loss of information when estimating the interaction term (Varela et al., 2006). A recent study combined location, management, and year into one factor named environment, such that the $G \times E$ can be assessed using the standard fixed-effect AMMI model (Paderewski et al., 2015).

As a method for combining information from several parallel data sources, hierarchical Bayesian modeling (Gelman, 2004) offers plant breeders the opportunity to use all available information collected throughout the trials established in sites and years. Bayesian modeling has further advantages over other approaches because it allows formal learning and borrowing of information from other experiments or from researchers' expertise through Bayes' rule. This paradigm also provides the flexibility of using data from previous studies. A Bayesian approach to study incomplete field trials, including data on genotypes, locations, and years, through a hierarchical model was presented by Theobald et al. (2002). Foucteau and Denis (2000) used Bayesian estimation to make variety recommendations using information from different multienvironment trials. Edwards and Jannink (2006) applied a Bayesian methodology for analyzing $G \times E$ with heterogeneous variance among environments. Cotes et al. (2006) described a Bayesian estimation for determining stable genotypes based on Shukla's stability variance (Shukla, 1972). Perez-Elizalde et al. (2012) listed the advantages of the conditional posterior estimates of Bayesian linear–bilinear models for plant breeding data: (i) they can be used with unequal cell size, (ii) they are an efficient test of the significance of the number of $G \times E$ bilinear terms, (iii) they identify genotypes and environments that may cause important $G \times E$, and (iv) they are efficient for incorporating information from historical plant breeding data. However, in this context, hierarchical Bayesian methods have not been investigated for studying adaptability patterns in multisite, multiyear plant breeding trials.

In this study, we propose extending the Bayesian model [2] to a hierarchical setting using the von Mises–Fisher distribution as the conditional conjugate prior distributions for the orthonormal matrices produced by the singular value decomposition of the interaction matrices. Bivariate highest probability density regions were estimated for the posterior distributions of the first two bilinear components of the superpopulation (across years) parameter as well as for each annual trial. We derive the posterior distributions of the first- and second-level parameters of the hierarchical Bayesian model and describe a Gibbs sampler of the joint posterior distributions. We illustrate the use of the hierarchical Bayesian model [2] by using both simulated and real plant breeding data. The simulated data set comprises four evaluation periods (years), while the real data set had 20 maize lines evaluated in 12 sites in three consecutive periods (years); in both cases, data from the first period is used to elicit the priors of the remaining periods.

MATERIALS AND METHODS

Simulated Data Set

Four periods representing 4 yr of evaluation comprising seven genotypes (1–7) and five sites (S1–S5) were simulated; the first evaluation period (year) was used as prior information for the remaining three periods. Each population was simulated according to a Gaussian model with mean $u + \beta_j + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk}$ (i = 1,2,…,7; j = 1,2,…,5) and variance $\sigma^2 = 1$. Three replicates were generated for each combination (cell) between the ith genotype and the jth site, yielding a total sample size of $7 \times 5 \times 3 = 105$ observations per period. Here we assume a complete randomized design of the seven genotypes in the five sites and 4 yr.

Experimental Data Set

The experimental data consisted of grain yield records (kg ha^{-1}) of 20 maize lines evaluated in 12 sites (two replicates per genotype × site combination) in three consecutive years. The experimental design used in each year × site combination was a randomized complete block design. For simplicity, models [1] and [2] expressed the average response of the genotypes $\overline{y}_{ij}$ (in terms of the values preadjusted by the design effects (replicates).

Statistical Model

The Hierarchical Bayesian Model

In the Bayesian model proposed by Perez-Elizalde et al. (2012), multivariate normal priors are assumed for the linear effects $\boldsymbol{\theta} = (\mu, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and the mVMF distributions for the orthonormal matrices ($\mathbf{U}$, $\mathbf{V}$), while each element of the diagonal matrix $\mathbf{D}$ follows a left-truncated normal distribution.

In this research, we set h as the total number of evaluation periods, while m denotes the mth evaluation period. Model [2] was used as the starting point for constructing a hierarchical model with a common mean, site linear effects given in $\boldsymbol{\theta}_m = (\mu_m, \boldsymbol{\beta}_m)$, and multiplicative terms given by the orthonormal matrices ($\mathbf{U}_m$, $\mathbf{V}_m$) and $\mathbf{D}_m$, whose priors were as previously stated. The proposed hierarchical model has two levels: the first level consists of experimental data from each evaluation period across sites, while the second level combines information from all periods (across periods) by including parameters of the population from which the observations supposedly came; thus, each period (year) may be considered as a realization of a superpopulation model. Note that the temporal correlation between years was not explicitly modeled; instead, between years correlations were implicitly considered in the hierarchical model.

The First-Level Model

Suppose there are h evaluation periods (years) with the mth evaluation period (m = 1,…,h) being in the form of model [2], such that

$$\mathbf{Y}_m = \mu_m \mathbf{1}_g \mathbf{1}'_s + \boldsymbol{\beta}'_m \otimes \mathbf{1}_g + \mathbf{U}_m \mathbf{D}_m \mathbf{V}'_m + \mathbf{E}_m \qquad [3]$$

Also consider the set of parameters $\{\mu_m, \beta_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{V}_m\}_{m=1}^{h}$ that index the distributions of periods $\mathbf{Y}_1$, $\mathbf{Y}_2$,…,$\mathbf{Y}_h$, forming exchangeable samples (i.e., exchangeable means that any permutation of any sample size of these random variables follows the same distribution; therefore, the order of the samples is not important). Details of the second-level model distribution are given below, after the likelihood function is defined.

Likelihood Function

The likelihood function for model [3] parameters in the mth period (year) is given by

$$\mathbf{L}\left(\boldsymbol{\theta}_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \tau_m \mid \mathbf{Y}_m\right) \propto \tau_m^{\frac{n_m gs}{2}} \exp\left\{-\frac{\tau_m}{2}\left[n_m \operatorname{tr}\left(\mathbf{E}_m \mathbf{E}'_m\right) - \left(n_m - 1\right)\operatorname{tr}\left(\mathbf{S}_m \mathbf{S}'_m\right)\right]\right\} \quad [4]$$

where $\tau_m = 1/\sigma_m^2$, $\mathbf{S}_m = \left\{\sqrt{s_{mij}^2}\right\}$, $s_{mij}^2 = \dfrac{\sum_{l=1}^{n}\left(\bar{Y}_{mij} - Y_{mijl}\right)^2}{n_{mij} - 1}$, $\mathbf{E}_m = \mathbf{Y}_m - \mu_m \mathbf{1}_g \mathbf{1}'_s - \boldsymbol{\beta}'_m \otimes \mathbf{1}_g - \mathbf{U}_m \mathbf{D}_m \mathbf{V}'_m$ and exp denotes the exponential function.

The matrix notation used in models [2] and [3] is convenient not only because it makes it easy to identify the interaction components, but also because it provides the flexibility to incorporate all available information to be used in further analyses (Perez-Elizalde et al., 2012). In this case, the use of the operator tr (in Eq. 4) denotes the trace of a matrix (i.e., the sum of the elements of the main diagonal of a square matrix). Also, under the classical (frequentist) approach, the expression in brackets in [4] is the residual sum of squares, which depends on the sum of squares as a result of the model minus (n_m − 1) times the sampling variance, with n_m as the number of replicates per cell in the mth period. The residual error variance for the mth period is denoted by σ_m^2; however, $\tau_m = \dfrac{1}{\sigma_m^2}$, the precision parameter (the inverse of the residual error variance), was used for convenience.

The conditional likelihood [6] (see below) derived from [4] was used to define conditional conjugate priors for the first level model parameters. As will be shown in Eq. [11–14], this criterion allows obtaining prior distributions that are easily elicited with information available in the form of cell prior averages and standard deviations.

The Second-Level Model

For the linear terms $\boldsymbol{\theta}_m = \left(\mu_m, \boldsymbol{\beta}_m\right)$ in model [3], we assumed a (1 + s) multivariate normal distribution with mean $\boldsymbol{\theta} = (\mu, \boldsymbol{\beta})$ and singular covariance matrix

$$\boldsymbol{\Sigma}_m = \left(n\tau_m\right)^{-1}\begin{bmatrix} (gs)^{-1} & \mathbf{0} \\ \mathbf{0} & (g)^{-1}\mathbf{K}_s\mathbf{K}'_s \end{bmatrix}$$

where $\mathbf{K}_s$ is a matrix of $s \times (s-1)$ such that $\mathbf{K}'_s\mathbf{K}_s = \mathbf{I}_{s-1}$ and $\mathbf{K}_s\mathbf{K}'_s = \mathbf{I}_s - \dfrac{1}{s}\mathbf{J}_s$, with $\mathbf{J}_s$ an $s \times s$ matrix with all elements equal to one; n is the prior cell size for the mth period. As a result of the restriction $\boldsymbol{\beta}'_m \mathbf{1}_g = \mathbf{0}$ established to avoid identifiability problems, a one-to-one transformation such as $\boldsymbol{\beta}_m^* = \mathbf{K}'_g \boldsymbol{\beta}_m$ is required. Let $\boldsymbol{\theta}_m^* = \left(\mu_m, \boldsymbol{\beta}_m^*\right)$; then its density function is

$$\pi\left(\theta_m^* \mid \Sigma_m^*\right) \propto \left|\Sigma_m^*\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\theta_m^* - \theta^*\right)'\Sigma_m^{*-1}\left(\theta_m^* - \theta^*\right)\right\}, \quad [5]$$

which is a multivariate normal distribution with mean $\boldsymbol{\theta}^*$ and a block diagonal covariance matrix given by $\boldsymbol{\Sigma}_m^* = \left(n\tau_m\right)^{-1}\begin{bmatrix} (gs)^{-1} & \mathbf{0} \\ \mathbf{0} & (g)^{-1}\mathbf{I}_{s-1} \end{bmatrix}$, where $\boldsymbol{\theta}^* = (\mu, \boldsymbol{\beta}^*)$ is the vector of means of the main effects.

As shown in Perez-Elizalde et al. (2012), given $\boldsymbol{\theta}_m$ and τ_m, the conditional likelihood function for the matrices ($\mathbf{U}_m$, $\mathbf{D}_m$, $\mathbf{V}_m$) is

$$\begin{aligned} &L\left(\mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m \mid \theta_m, \tau_m, \mathbf{Y}_m\right) = \\ &L\left(\mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m \mid \tau_m, \mathbf{Y}_m\right) \propto \\ &\exp\left\{-\frac{n_m\tau_m}{2}\operatorname{tr}\left[\left(-2\mathbf{Y}_m + \mathbf{U}_m\mathbf{D}_m\mathbf{V}'_m\right)\left(\mathbf{U}_m\mathbf{D}_m\mathbf{V}'_m\right)'\right]\right\} = \\ &\operatorname{etr}\left\{-\frac{n_m\tau_m}{2}\left(-2\mathbf{Y}_m + \mathbf{U}_m\mathbf{D}_m\mathbf{V}'_m\right)\left(\mathbf{U}_m\mathbf{D}_m\mathbf{V}'_m\right)'\right\} \end{aligned} \quad [6]$$

where etr denotes the exponential of the trace function. Thus, mathematically convenient conditional distributions for $\mathbf{U}_m$ and $\mathbf{V}_m$ are in the form

$$\pi\left(\mathbf{U}_m \mid \mathbf{V}, \mathbf{D}, \tau_m, \mathbf{M}_{0m}\right) \propto \operatorname{etr}\left(n\tau_m\mathbf{M}_{0m}\mathbf{V}\mathbf{D}\mathbf{U}'_m\right) \quad [7]$$

and

$$\pi\left(\mathbf{V}_m \mid \mathbf{U}, \mathbf{D}, \tau_m, \mathbf{M}_{0m}\right) \propto \operatorname{etr}\left(n\tau_m\mathbf{M}'_{0m}\mathbf{U}\mathbf{D}\mathbf{V}'_m\right) \quad [8]$$

which are mVMF distributions, with $\mathbf{M}_{0m}$ as the prior predicted cell matrix for the mth evaluation period.

For easy handling, each element on the diagonal of $\mathbf{D}_m$, $\lambda_{1_m} > \lambda_{2_m} > \ldots > \lambda_{t_m}$ was considered a conditional conjugate left-truncated normal distribution with density given by

$$\pi\left(\lambda_{k_m} \mid \tau_m, l_k\right) = \left\{1 - \phi\left(\sqrt{n\tau_m}\left(\lambda_{(k+1)_m} - l_k\right)\right)\right\}^{-1} \mathrm{N}\left(\lambda_{k_m} \mid l_k, \left(n\tau_m\right)^{-1}\right) \quad [9]$$

where l_1, l_2,…,l_t are the diagonal elements of $\mathbf{D}$, and N(.) and ϕ denote the probability density function and the cumulative distribution function of a normal distribution, respectively.

For the precision parameter τ_m, a conjugate gamma distribution with parameters $a/2$ and $S_0^2/2$ was stated by

$$\pi(\tau_m) \propto \tau_m^{(a/2-1)} \exp\left\{\frac{-(aS_0^2)}{2}\tau_m\right\} \quad [10]$$

The Prior Distribution

To complete the hierarchical model, prior distributions were assigned to the superpopulation parameters (μ, $\boldsymbol{\beta}$, **D**, **U**, **V**). For linear terms $\boldsymbol{\theta}^* = (\mu, \boldsymbol{\beta}^*)$, the conjugate prior density is

$$\pi\left(\boldsymbol{\theta}^* \mid \Sigma_p^*\right) \propto \left|\Sigma_p^*\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0^*\right)' \Sigma_p^{*-1}\left(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0^*\right)\right\} \quad [11]$$

where $\boldsymbol{\Sigma}_p^* = n_0\tau_0 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & (g)^{-1}\mathbf{I}_{s-1} \end{bmatrix}$ and $\boldsymbol{\theta}^* = \left(\mu_0, \mathbf{K}_s'\boldsymbol{\beta}_0\right)$.

The conjugate priors for **U** and **V** are mVMF distributions with densities given by

$$\pi\left(\mathbf{U} \mid \mathbf{V}_0, \mathbf{D}_0, \tau_0, \mathbf{M}_0\right) \propto \operatorname{etr}\left(n_0\tau_0\mathbf{M}_0\mathbf{V}_0\mathbf{U}'\right) \quad [12]$$

and

$$\pi\left(\mathbf{V} \mid \mathbf{U}_0, \mathbf{D}_0, \tau_0, \mathbf{M}_0\right) \propto \operatorname{etr}\left(n_0\tau_0\mathbf{M}'_0\mathbf{U}_0\mathbf{V}'\right) \quad [13]$$

In the above priors, available information can be introduced through corresponding hyperparameters. For example, $\mathbf{U}_0$, $\mathbf{D}_0$, and $\mathbf{V}_0$ could be seen as the singular value decomposition of $\mathbf{Z}_0 = \mathbf{M}_0 - \mu_0\mathbf{1}_g\mathbf{1}'_s - \boldsymbol{\beta}'_0 \otimes \mathbf{1}_g$, where $\mathbf{M}_0$ is the matrix of prior predicted cell means with n_0 elements per cell; μ_0 and $\boldsymbol{\beta}_0$ are prior beliefs about the general mean and column effects, respectively.

For the elements of the diagonal matrix **D** of singular values, we used conjugate priors left-truncated normal distributions with density

$$\pi\left(\lambda_k \mid \tau_0\right) = \left\{1 - \Phi\left(\sqrt{n_0\tau_0}\left(\lambda_{k+1} - l_k^0\right)\right)\right\}^{-1} \mathrm{N}\left(\lambda_k \mid l_k^0, n_0\tau_0^{-1}\right) \quad [14]$$

with $k = 1,2,\ldots,t$ and $\lambda_{t+1} = 0$. In the above equation $\left(l_1^0, l_2^0, \ldots, l_t^0\right) = \operatorname{diag}\left(\mathbf{U}'_0\mathbf{Y}_0\mathbf{V}_0\right)$, the hyperparameter l_k^0 is the prior mean of the kth singular value, which may be elucidated through the previously described singular value decomposition; τ_0 and n_0 are the estimated precision parameter and the prior size of the element in $\mathbf{M}_0$, respectively.

The Posterior Distribution

The joint posterior distribution is obtained following Bayes' rule as the product between the evaluation period likelihoods in [4], the joint second level distributions given by [5] and [7–10], and the joint prior obtained as the product of densities [11–14]. So, the joint posterior is given by

$$\prod_{m=1}^{h} L\left(\boldsymbol{\theta}_m^*, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \tau_m \mid \mathbf{Y}_m\right) \times \pi\left(\boldsymbol{\theta}_m^*, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \tau_m \mid \boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}\right)\pi\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}\right) \quad [15]$$

Since the joint posterior [15] is obviously high dimensional, an MCMC approach is necessary to obtain estimations of marginal posterior distributions and summaries. The Bayesian implementation can be entirely on a Gibbs sampler because, under stated prior assumptions, the fully conditional posterior distributions have a closed form. A brief description of the fully conditional distributions required for Bayesian implementation is given in Appendix A. In practice, the Gibbs sampler can be implemented by sampling iteratively from the full conditional posterior distributions; then, after a number of iterations large enough to reach convergence on the joint posterior distribution, an MCMC sample is obtained and used to calculate summaries such as means, standard deviations, and HPD intervals.

The Gibbs Sampler

The Gibbs sampler comprises a total of W iterations; for the wth iteration $\{w = 1,\ldots,W\}$, samples from the parameters of each period (first stage), and samples from the superpopulation (all periods) parameters (second stage) are sequentially drawn. This implies that in the wth iteration, samples from the set of first level posterior distributions are obtained for each period $\{m = 1,\ldots,h\}$ and once samples for all parameters of all periods are obtained, samples for the set of the second level posterior distributions (superpopulation parameters) are drawn (second stage). Thus, an iteration is completed after obtaining samples for all parameters in the first and second stages.

The next set of posterior distributions is used to draw samples for the mth $\{m = 1,\ldots,h\}$ period of the first stage. This sequence is repeated until samples from all periods have been drawn.

$$\tau_m^{(w)} \sim \pi\left(\tau_m \mid a_m, b_m, \mathbf{E}_m\right)$$

$$\boldsymbol{\theta}_m^{*(w)} \sim \mathrm{N}_c\left\{\boldsymbol{\theta}_m^* \mid \left(\Sigma_{am}^{*-1} + \Sigma_m^{*-1}\right)^{-1}\left(\Sigma_{am}^{*-1}\hat{\boldsymbol{\theta}}_m^* + \Sigma_m^{*-1}\boldsymbol{\theta}^*\right), \left(\Sigma_{am}^{*-1} + \Sigma_m^{*-1}\right)^{-1}\right\}$$

$$\mathbf{D}_m^{(w)} \sim \pi\left(\mathbf{D}_m \mid \tau_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{Y}_m, \mathbf{D}\right) \text{ (see Appendix A)}$$

$$\mathbf{U}_m^{(w)} \sim \pi\left(\mathbf{U}_m \mid \tau_m, \mathbf{V}_m, \mathbf{D}_m, \mathbf{Y}_m, \mathbf{M}_{0m}, \mathbf{V}, \mathbf{D}\right) \propto \operatorname{etr}\left(\tau_m\left[n_m\mathbf{Y}_m\mathbf{V}_m\mathbf{D}_m + n\mathbf{M}_{0m}\mathbf{V}\mathbf{D}\right]\mathbf{U}'_m\right)$$

$$\mathbf{V}_m^{(w)} \sim \pi\left(\mathbf{V}_m \mid \tau_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{Y}_m, \mathbf{M}_{0m}, \mathbf{U}, \mathbf{D}\right) \propto \operatorname{etr}\left(\tau_m\left[n_m\mathbf{Y}'_m\mathbf{U}_m\mathbf{D}_m + n\mathbf{M}'_{0m}\mathbf{U}\mathbf{D}\right]\mathbf{V}'_m\right)$$

After drawing samples for each set of parameters of each period, the superpopulation parameters are computed in the second stage

$$\boldsymbol{\theta}^{*(w)} \sim \mathrm{N}_c\left\{\boldsymbol{\theta}^* \mid \left(\sum_{m=1}^{h}\Sigma_m^{*-1} + \Sigma_p^{*-1}\right)^{-1}\left(\sum_{m=1}^{h}\Sigma_m^{*-1}\boldsymbol{\theta}_m^* + \Sigma_p^{*-1}\boldsymbol{\theta}_0^*\right), \left(\sum_{m=1}^{h}\Sigma_m^{*-1} + \Sigma_p^{*-1}\right)^{-1}\right\}$$

$$\mathbf{D}^{(w)} \sim \pi\left(\mathbf{D} \mid \tau_m, \mathbf{D}_m\right)$$

$$\mathbf{U}^{(w)} \sim \pi\left(\mathbf{U} \mid \tau_m,, \mathbf{V}_m, \mathbf{D}_m, \mathbf{M}_{0m}\right) \propto \operatorname{etr}\left(\left[n\sum_{m=1}^{h}\tau_m\mathbf{M}_{0m}\mathbf{V}\mathbf{D} + n_0\tau_0\mathbf{M}_0\mathbf{V}_0\mathbf{D}_0\right]\mathbf{U}'\right)$$

$$\mathbf{V}^{(w)} \sim \pi\left(\mathbf{V} \mid \tau_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{M}_{0m}\right)$$
$$\propto \text{etr}\left(\left[\sum_{m=1}^{h} \tau_m \mathbf{M}'_{0m} \mathbf{UD} + n_0 \tau_0 \mathbf{M}'_0 \mathbf{U}_0 \mathbf{D}_0\right] \mathbf{V}'\right)$$

Details of the listed posterior distributions used to draw samples for the first and second stages can be found in Appendix A. Posterior distributions were computed using two parallel chains of size 35,000 with a burn-in period of 15,000 and a thinning interval of 4. Therefore, the size of the final MCMC was 10,000 iterations, which were used to compute statistical summaries about the posterior distribution. The Raftery and Lewis (1995) and Gelman and Rubin (1992) tests were used to assess convergence, in addition to visual inspection of the traces and autocorrelation function (ACF) plots (data not shown). Computational time to simulate each MCMC entirely using a conventional computer (x64 Pentium 4 Xeon 3.4 GHz, 4 GB RAM) was close to 60 min.

Availability of Simulated Data, Real Data, and R Codes for the Hierarchical Bayesian Model

Zip folders and files with real and simulated data, R codes, and supplemental figures are located in the following repository: http://hdl.handle.net/11529/10463. The folder with the real data has three Excel files, YEAR_1, YEAR_2, and YEAR_3, whereas the folder with the simulated data has four Excel files, period_1, period_2, period_3, and period_4. The R codes are in the folder HB.SREG that contains six subfolders, codes, data, docs, functions, input, and output. The supplemental figures containing results from the standard SREG biplot for each year (Year 1–3) with genotypes and sites are in Supplemental Fig. S1a–c and Supplemental Fig. S2a–c, respectively.

RESULTS AND DISCUSSION

Simulated Data Set

A simulation study was conducted to illustrate the performance of the proposed hierarchical Bayesian version of model [2]. Parameters of linear terms and variance were $\mu = 5$, $\beta = (-0.83 \;\; 0.50 \;\; -0.83 \;\; 0.33 \;\; 0.83)$ and $\sigma^2 = 1$, respectively. The bilinear components derived from singular value decomposition were:

$$\mathbf{U} = \begin{bmatrix} -0.23 & 0.41 & -0.12 & -0.47 \\ -0.02 & -0.04 & 0.00 & 0.54 \\ 0.49 & -0.46 & 0.49 & -0.16 \\ 0.77 & 0.40 & -0.29 & 0.23 \\ 0.05 & -0.50 & -0.16 & -0.26 \\ 0.01 & 0.42 & 0.78 & -0.05 \\ -0.31 & -0.10 & 0.14 & 0.56 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} 2.42 & 0 & 0 & 0 \\ 0 & 1.88 & 0 & 0 \\ 0 & 0 & 0.80 & 0 \\ 0 & 0 & 0 & 0.53 \end{bmatrix}, \text{ and}$$

$$\mathbf{V} = \begin{bmatrix} 0.65 & -0.52 & -0.39 & 0.31 \\ -0.03 & 0.32 & -0.76 & 0.00 \\ -0.49 & -0.28 & -0.48 & -0.38 \\ 0.48 & 0.64 & -0.08 & -0.33 \\ -0.29 & 0.33 & -0.10 & 0.80 \end{bmatrix}$$

Therefore, the matrix that accounts for the main effect of genotypes (G) plus $G \times E$ is given by $\mathbf{Z}$:

$$\mathbf{Z} = \begin{bmatrix} -0.802 & 0.336 & 0.198 & 0.316 & 0.226 \\ 0.096 & -0.023 & -0.064 & -0.166 & 0.218 \\ 1.041 & -0.610 & -0.495 & 0.012 & -0.736 \\ 0.948 & 0.361 & -1.059 & 1.354 & -0.172 \\ 0.575 & -0.207 & 0.318 & -0.488 & -0.443 \\ -0.646 & -0.222 & -0.522 & 0.476 & 0.170 \\ -0.342 & -0.123 & 0.254 & -0.587 & 0.382 \end{bmatrix}$$

As previously described, the proposed model allows incorporating available prior information at two levels; for individual periods (analyzed in the first level), prior information was supplied with estimates computed using the iterative least squares method on the first period ($\mathbf{M}_{0m}$), while for the second level, which combines information from all periods, no prior information was included at the superpopulation level ($\mathbf{M}_0$) by setting $\tau_0 = 0$, such that the posterior distribution of each parameter was the result of the weighted value of all periods.

Focusing on bilinear parameters, the resulting biplots of the first two components showed interaction effects between genotypes and sites (data not shown). Table 1 shows statistical summaries of the superpopulation parameters. The estimated eigenvalues were close to the true values; also, the second estimated eigenvector values were roughly two-thirds of the first component with larger standard deviation. The elements of the first two singular vectors that do not include the null point (0,0) for the bivariate 0.95 HPD region causing a sizable interaction (Crossa et al., 2011; Perez-Elizalde et al., 2012) were $u_{1,1}$, $u_{2,1}$, $u_{3,1}$, $u_{4,2}$, $u_{6,2}$, $u_{7,2}$, $v_{1,1}$, $v_{2,1}$, $v_{3,1}$, and $v_{4,2}$.

For easy understanding, only the relevant figures derived from the superpopulation-level analysis were depicted, since relevant figures for individual periods were similar (data not shown). The HPD regions of paired eigenvector elements that were statistically different from the null point (0,0) are depicted in Fig. 1a–b. Figure 1a shows the plot of the genotype scores $\text{UD}^{1/2}$, where the outer and inner shaded regions denote the bivariate 0.99 and

Table 1. Simulated data. For the superpopulation, the posterior summary (mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.95 highest posterior density (HPD), and 0.99 HPD intervals were computed using 10,000 approximately independent samples from the joint posterior distribution for linear effects (μ, $\beta_1 - \beta_2$), all the singular values ($\lambda_1 - \lambda_4$), and the first two left and right singular vector elements of rows (u_{ik}) and columns (v_{ik}), respectively.

Parameter	True value	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	0.95 HPD interval		0.99 HPD interval	
							Lower	Upper	Lower	Upper
μ	5	5.097	0.056	5.059	5.097	5.134	5.004	5.187	4.988	5.206
β_1	−0.83	−0.716	0.113	−0.793	−0.716	−0.641	−0.899	−0.528	−0.934	−0.49
β_2	0.5	0.641	0.113	0.565	0.642	0.717	0.457	0.827	0.421	0.861
β_3	−0.83	−0.946	0.112	−1.021	−0.945	−0.87	−1.131	−0.765	−1.166	−0.726
β_4	0.33	0.289	0.113	0.212	0.289	0.365	0.105	0.476	0.064	0.505
β_5	0.83	0.732	0.112	0.656	0.732	0.808	0.549	0.916	0.508	0.944
$u_{1,1}$	−0.23	−0.387	0.141	−0.482	−0.401	−0.31	−0.614	−0.173	−0.648	−0.102
$u_{2,1}$	−0.02	−0.343	0.13	−0.431	−0.354	−0.269	−0.554	−0.145	−0.586	−0.085
$u_{3,1}$	0.49	0.612	0.166	0.545	0.646	0.722	0.394	0.841	0.298	0.865
$u_{4,1}$	0.77	−0.033	0.204	−0.166	−0.034	0.098	−0.367	0.309	−0.444	0.375
$u_{5,1}$	0.05	0.13	0.143	0.041	0.135	0.223	−0.109	0.352	−0.144	0.419
$u_{6,1}$	0.01	0.249	0.222	0.123	0.27	0.401	−0.096	0.608	−0.183	0.685
$u_{7,1}$	−0.31	−0.228	0.225	−0.382	−0.244	−0.095	−0.606	0.115	−0.67	0.21
$u_{1,2}$	0.41	−0.028	0.252	−0.198	−0.025	0.141	−0.451	0.388	−0.524	0.465
$u_{2,2}$	−0.04	0.117	0.21	−0.015	0.129	0.263	−0.218	0.467	−0.305	0.524
$u_{3,2}$	−0.46	−0.254	0.273	−0.449	−0.278	−0.095	−0.717	0.149	−0.762	0.283
$u_{4,2}$	0.4	0.394	0.217	0.28	0.426	0.543	0.054	0.719	−0.046	0.776
$u_{5,2}$	−0.5	−0.172	0.2	−0.305	−0.184	−0.052	−0.495	0.152	−0.561	0.233
$u_{6,2}$	0.42	0.393	0.242	0.256	0.426	0.565	0.027	0.777	−0.083	0.823
$u_{7,2}$	−0.1	−0.451	0.226	−0.61	−0.487	−0.336	−0.797	−0.124	−0.826	0.01
$v_{1,1}$	0.65	0.501	0.157	0.425	0.524	0.606	0.272	0.742	0.185	0.775
$v_{2,1}$	−0.03	−0.664	0.151	−0.762	−0.693	−0.604	−0.873	−0.457	−0.906	−0.374
$v_{3,1}$	−0.49	−0.254	0.21	−0.399	−0.266	−0.127	−0.601	0.077	−0.653	0.166
$v_{4,1}$	0.48	0.135	0.273	−0.039	0.145	0.322	−0.305	0.588	−0.414	0.673
$v_{5,1}$	−0.29	−0.033	0.242	−0.191	−0.033	0.123	−0.417	0.381	−0.505	0.454
$v_{1,2}$	−0.52	−0.182	0.255	−0.361	−0.199	−0.025	−0.598	0.224	−0.666	0.323
$v_{2,2}$	0.32	0.137	0.295	−0.049	0.148	0.341	−0.322	0.647	−0.438	0.723
$v_{3,2}$	−0.28	−0.368	0.238	−0.531	−0.39	−0.233	−0.756	−0.003	−0.807	0.112
$v_{4,2}$	0.64	0.578	0.21	0.479	0.615	0.722	0.276	0.891	0.144	0.92
$v_{5,2}$	0.33	0.358	0.313	0.184	0.396	0.579	−0.113	0.865	−0.259	0.922
λ_1	2.42	1.874	0.401	1.611	1.888	2.148	1.218	2.54	1.06	2.629
λ_2	1.88	1.218	0.352	0.986	1.229	1.461	0.651	1.805	0.52	1.892
λ_3	0.8	0.751	0.326	0.519	0.754	0.977	0.203	1.287	0.089	1.343
λ_4	0.53	0.353	0.23	0.169	0.325	0.505	0	0.672	0	0.773

0.95 HPD posterior regions, respectively. Only Genotype 3, located in the lower right quadrant of the biplot, was different from the others, causing interaction at the 0.95 HPD. On the other hand, Fig. 1b shows the HPD regions for the sites by considering $VD^{1/2}$. The corresponding 0.95 HPDs of sites S1, S2, and S4 (located in three different quadrants) did not include the null point (0,0); therefore, they are considered the sites that cause most interaction. Across periods, the proposed model allows inferring that Genotype 3 has the best rank in site S1 (1.01) (same direction, same quadrant), but it is the worst-ranked genotype in site S2 (0.61), as they (Genotype 3 and site S2) are located in the opposite direction of the biplot.

Histograms derived from sampling the marginal posterior densities of eigenvalues ($\lambda_1 - \lambda_4$) are depicted in Fig. 1c. The posterior densities of late singular values move quickly toward zero, as expected. Figure 1d shows the posterior densities of the cumulative variance; the first component explains roughly 76% of the interaction variance, and the first two components captured nearly 93% of this variability.

In general, the estimated parameters within the estimated intervals were close to the true values. Minor discrepancies between parameters (signal) and estimations at the superpopulation level were found, which may be due to (i) noise, (ii) induced shrinkage, or (iii) site effects. Results obtained with the proposed models showed that the hierarchical Bayesian approach offers good inferential answers to two-way table analysis when jointly estimating the effects of genotypes and interactions (genotype × site) within and across periods. Results of Periods 2, 3, and 4 follow similar patterns when comparing true and estimated parameters (data not shown). The empirical

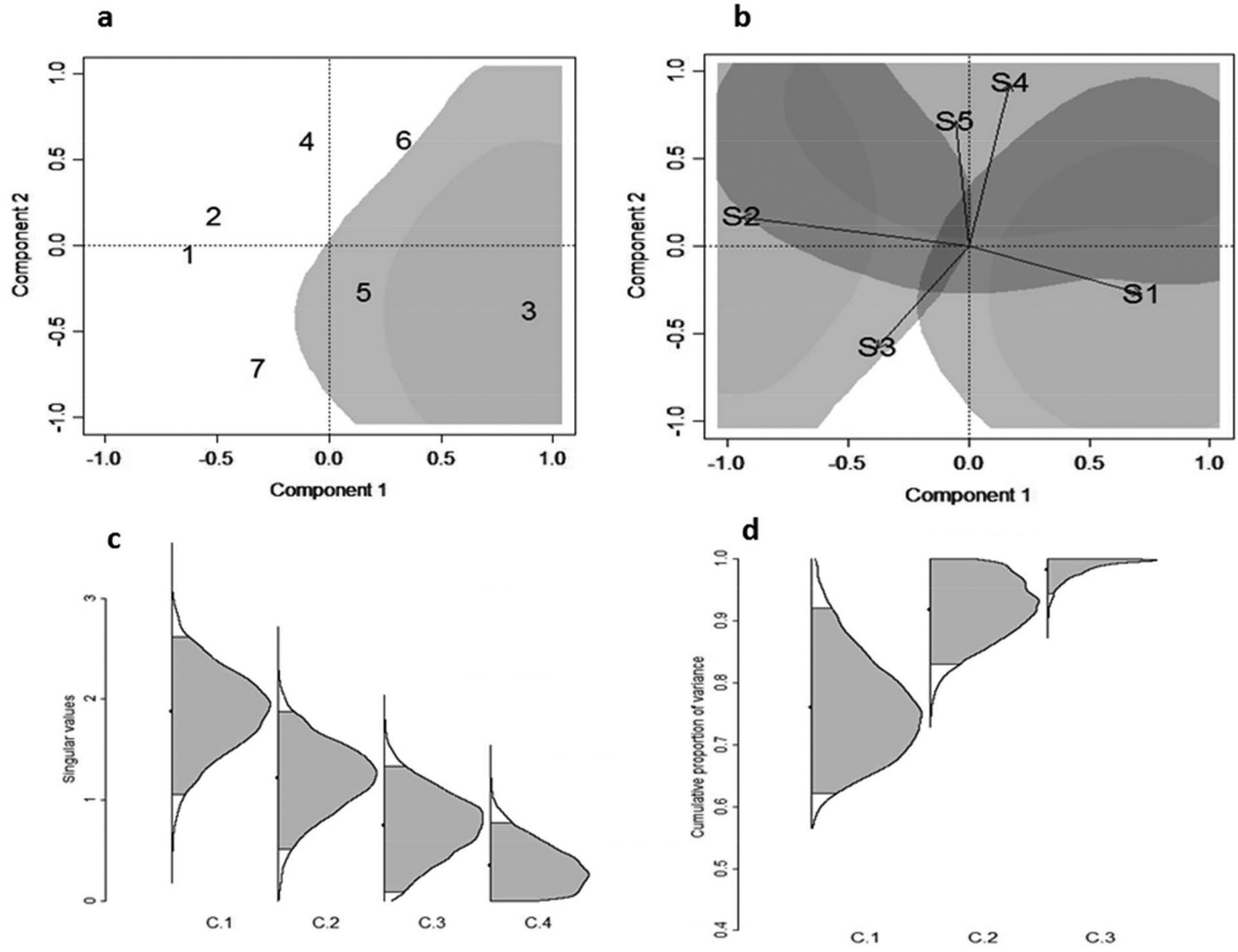

Fig. 1. Simulated data. Superpopulation estimated multiplicative parameters derived by analyzing three periods (Years 2, 3, and 4), each containing seven rows (1–7) and five columns (S1–S7): (a) plot of the genotype scores $UD^{1/2}$ and the bivariate 0.99 (gray external contour) and 0.95 (gray internal contour) highest posterior density (HPD) regions; (b) plot of the site scores $VD^{1/2}$ and corresponding bivariate 0.99 and 0.95 HPD regions. Only Genotype 3 and Sites S1, S2 and S4 did not include the null point at the 0.95 HPD probability level; (c) posterior densities and 0.95 HPD regions of the singular values, (λ_1,…,λ_4 (C.1–C.4); (d) posterior densities and 0.95 HPD regions of the cumulative proportion of variance $\varphi_t = \frac{\sum_{k=1}^{t}\lambda_k^2}{\sum_{k=1}^{\min(r,c)-1}\lambda_k^2}$, t = 1,..,min(r,c) – 2.

results of the simulation using the proposed hierarchical Bayesian version of model [2] (which allows borrowing information among periods through superpopulation parameters) showed that it might give reliable estimations of the remaining unknowns in the model.

Experimental Data Set

Superpopulation Analyses

The HPD regions of genotypic and environmental scores that are very likely to be different from the null point (0,0) are shown in Table 2 and in the biplots in Fig. 2a and 3a, for genotypes and sites, respectively. Table 2 gives the posterior means of genotypic and site scores that do not include the null point (0,0) in their 0.90 HPD regions; 0.95 HPD intervals are also presented. Figure 2a shows the plot of the posterior genotypic scores $UD^{1/2}$, where the outer and inner shaded areas of the graph are the bivariate 0.95 and 0.90 HPD posterior regions, respectively. Only Genotypes 1, 2, and 5 were depicted because their corresponding 0.90 HDP regions did not include the null point (0,0). Similarly, Fig. 3a shows the biplot of the posterior site scores $VD^{1/2}$ for environments (S8 and S11) that did not include the null point (0,0) in their 0.90 HPD regions.

Since there is a clear overlapping between the interaction scores of Genotypes 1 and 2 at the 0.90 HPD regions (Fig. 2a), these two genotypes did not produce statistical differences in the $G \times E$ effects. On the other hand, since there is no overlapping between the scores of Genotype 5 and Genotypes 1 and 2 at the 0.90 HPD, these two groups of genotypes produced sizable $G \times E$. Similarly,

Table 2. Experimental data. For the superpopulation, the posterior summary (mean), standard deviation (SD), quartile ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.90 highest posterior density (HPD), and 0.95 HPD intervals were computed using 10,000 approximately independent samples simulated from the joint posterior distribution for singular values ($\lambda_1,\ldots,\lambda_{11}$) and the left and right singular vector elements of genotypes (u_{ik}) and sites (v_{jk}), respectively, whose 0.90 HPD and 0.95 HPD intervals do not contain the null value (0,0). Grain yield data were measured in kilograms per hectare.

						0.90 HPD interval		0.95 HPD interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	Lower	Upper	Lower	Upper
λ_1	2373.1	510.6	2100.1	2421.6	2714.3	1604.7	3197.9	1329.1	3324.4
λ_2	943.0	469.5	593.9	922.7	1268.8	145.2	1678.3	28.7	1772.9
λ_3	434.3	303.5	192.0	386.6	625.6	0.0	860.8	0.0	1001.7
λ_4	201.2	185.0	55.5	150.1	295.9	0.0	459.8	0.0	573.0
λ_5	95.0	108.3	16.5	55.9	137.0	0.0	243.2	0.0	317.7
λ_6	45.5	63.3	4.9	20.1	60.3	0.0	126.5	0.0	177.2
λ_7	22.1	36.7	1.5	7.4	26.4	0.0	63.0	0.0	95.3
λ_8	10.9	21.3	0.5	2.7	11.2	0.0	31.4	0.0	50.5
λ_9	5.4	12.2	0.2	1.0	4.8	0.0	14.8	0.0	26.4
λ_{10}	2.7	7.2	0.0	0.4	2.0	0.0	7.0	0.0	13.2
λ_{11}	1.3	4.2	0.0	0.1	0.8	0.0	3.2	0.0	6.5
$u_{1,1}$	−0.460	0.145	−0.555	−0.487	−0.400	−0.676	−0.259	−0.700	−0.163
$u_{2,1}$	−0.387	0.157	−0.489	−0.413	−0.316	−0.626	−0.163	−0.653	−0.059
$u_{5,1}$	0.332	0.140	0.257	0.346	0.425	0.125	0.562	0.044	0.588
$u_{14,1}$	0.187	0.117	0.117	0.193	0.264	0.005	0.380	−0.043	0.419
$u_{17,1}$	0.224	0.123	0.153	0.233	0.305	0.031	0.425	−0.022	0.462
$v_{5,1}$	−0.205	0.131	−0.289	−0.211	−0.129	−0.417	−0.004	−0.456	0.052
$v_{6,1}$	0.275	0.137	0.202	0.287	0.365	0.060	0.489	0.004	0.535
$v_{8,1}$	−0.337	0.123	−0.417	−0.348	−0.271	−0.531	−0.157	−0.563	−0.099
$v_{11,1}$	−0.591	0.168	−0.695	−0.620	−0.530	−0.821	−0.395	−0.853	−0.284

the environmental scores for S8 and S11 (Table 2; Fig. 3a) did not include the null point (0,0) at the 0.90 HPD, and provided enough evidence to claim that $G \times E$ effects induced in these environments were significant.

The response of genotypes, sites, and the joint response of genotypes and sites can be examined by considering Fig. 2a and 3a simultaneously. Genotype 5 formed a group by itself, while Genotypes 1 and 2 formed a distinct group that shows differential responses in different environments. Since Genotypes 1 and 2 point in a similar direction (left-hand side of Fig. 2a) as environments S8 and S11 (left-hand side of Fig. 3a), we concluded that they had positive genotypic main effects and $G \times E$ interactions at those sites, whereas Genotype 5, which is located in a direction (lower right-hand quadrant of Fig. 3a) opposite to sites S8 and S11, had negative genotypic main effects and $G \times E$ interaction in sites S8 and S11. Sample summaries (means, standard deviations, quartiles, etc.) of the linear effects for superpopulation analysis are given in Table B1 of Appendix B.

In summary, the hierarchical Bayesian analysis found patterns of genotypic main effects plus $G \times E$ between genotypes and sites; it provided useful information using the joint response of genotypes and sites from several years. Meta-analysis using the Bayesian hierarchical approach for identifying subsets of homogenous genotypes and sites that cause significant interaction is useful to breeders because all available information is included in the estimation of superpopulation parameters. The values of these parameters can be viewed as the result of combining several evaluation periods, allowing the information contained in those periods to show the general patterns of the main genotypic and $G \times E$ effects. Regarding periodic evaluations, it is possible to find main genotypic and $G \times E$ effects that are specific to the current evaluation period; for this reason, it is important to consider the analysis of each evaluation period to observe changes from one period to another as a result of a year effect.

Yearly Analyses

From a plant breeding perspective, assessing the overall adaptation of genotypes to environmental conditions through several evaluation periods is important, as is assessing their specific adaptation within each evaluation period. When Years 2 and 3 were analyzed taking into account superpopulation parameters, the genotypes and sites that contributed most to the interaction were found. For example, Genotypes 1 and 2 and sites S2, S5, S8, and S11 showed genotypic and $G \times E$ effects when Years 2 and 3 were analyzed (Table 3, 4; Fig. 2b–c, 3b–c); these genotypes and sites were also important at the superpopulation level. Furthermore, biplots from the different evaluation periods provided information on specific responses that depend mainly on the current evaluation period (year effect).

In Year 2, only Genotype 1 showed genotypic and $G \times E$ effects that do not include the bivariate null point (0,0) (Fig. 2b); whereas, in Year 3, Genotypes 1 and 2 had

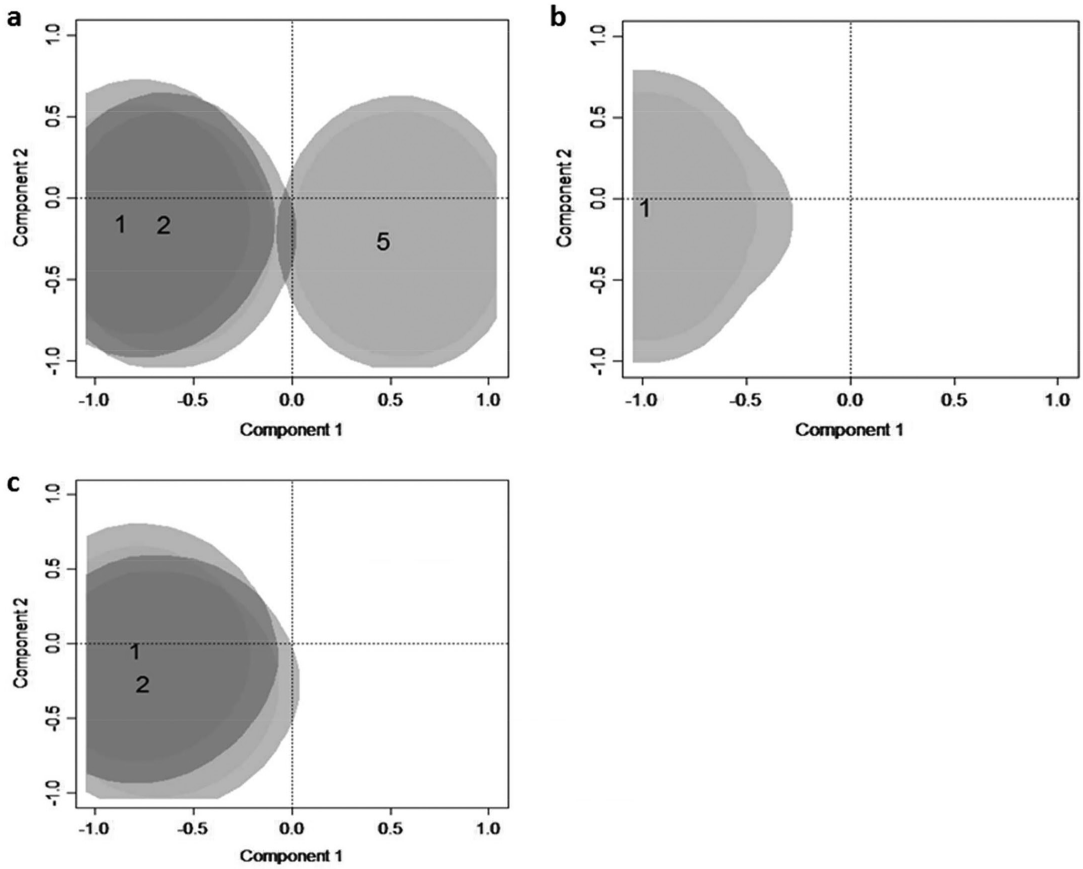

Fig. 2. Experimental data of 20 genotypes in 12 sites; plot of the row (genotypes) scores $UD^{1/2}$ and the bivariate 95% (gray external contour) and 90% (gray internal contour) highest posterior density (HPD) regions (only genotypes that do not include the null point (0,0) at the 0.90 HPD probability level are depicted) for: (a) superpopulation, Genotypes 1, 2, and 5; (b) analysis of the second evaluation period (Year 2), Genotype 1; (c) analysis of the third evaluation period (Year 3), Genotypes 1 and 2.

genotypic and $G \times E$ effects that do not include the null point (0,0) (Fig. 2c). As for sites, the 0.90 and 0.95 HPD regions of sites S2, S5, S8, and S11 overlapped and formed a group of sites with sizeable interaction (Fig. 3b). A similar pattern occurred in Year 3, when Genotypes 1 and 2 formed one main group (Fig. 2c), and sites that did not include the null bivariate point (0,0) formed another main group (S8, S11) (Fig. 3c). Regarding genotypes, results of yearly analyses follow similar patterns as those obtained with the superpopulation model, that is, Genotypes 1 and 2 being the ones causing $G \times E$; the only difference is that Genotype 5 did not show genotypic and $G \times E$ effects in any year. These facts indicate that analyses of the superpopulation captured the complexity of year-to-year variation between genotypes and sites in a parsimonious model.

Posterior densities of eigenvalues for the superpopulation and Periods 2 and 3 are depicted in Fig. 4a–c, respectively. In all three cases, the densities collapsed at zero beyond the third component. The posterior densities of the cumulative proportion of variance are shown in Fig. 5a–c for the superpopulation and Periods 2 and 3, respectively. Results showed that in all three cases the first component explains roughly 95% of the genotype plus interaction variance, while the first two components capture nearly 98% of this variability. The sample means of linear effects for Periods 2 and 3 are given in Tables B2 and B3 in Appendix B, respectively.

Results from the Standard Model versus the Hierarchical Bayesian Model

Although biplots obtained from the standard ordinary least square fit (SREG) and the hierarchical Bayesian model proposed in this study are not strictly comparable, we give a brief comparison of results. Supplemental Fig. S1a shows the results of the standard biplot for the first year, and Fig. 2a depicts the superpopulation biplot including Year 1 as the prior information. Both biplots show Genotypes 1 and 2 vs. Genotype 5 in different opposite quadrants as those contributing the most to the $G \times E$. Also, the contribution to the $G \times E$ of Genotypes 1 and 2 is clear for the Year 1 analyses from the hierarchical Bayesian model (Fig. 2a) but not from the standard SREG biplot (Supplemental Fig. S1b), where Genotype 2 is located toward the center of the biplot. Similarities and differences can also be examined when comparing the distribution of sites in the yearly analyses from the standard biplot (Supplemental Fig. S2a–c) as compared with those obtained from the hierarchical Bayesian model (Fig. 3a–c). Site S11 appears

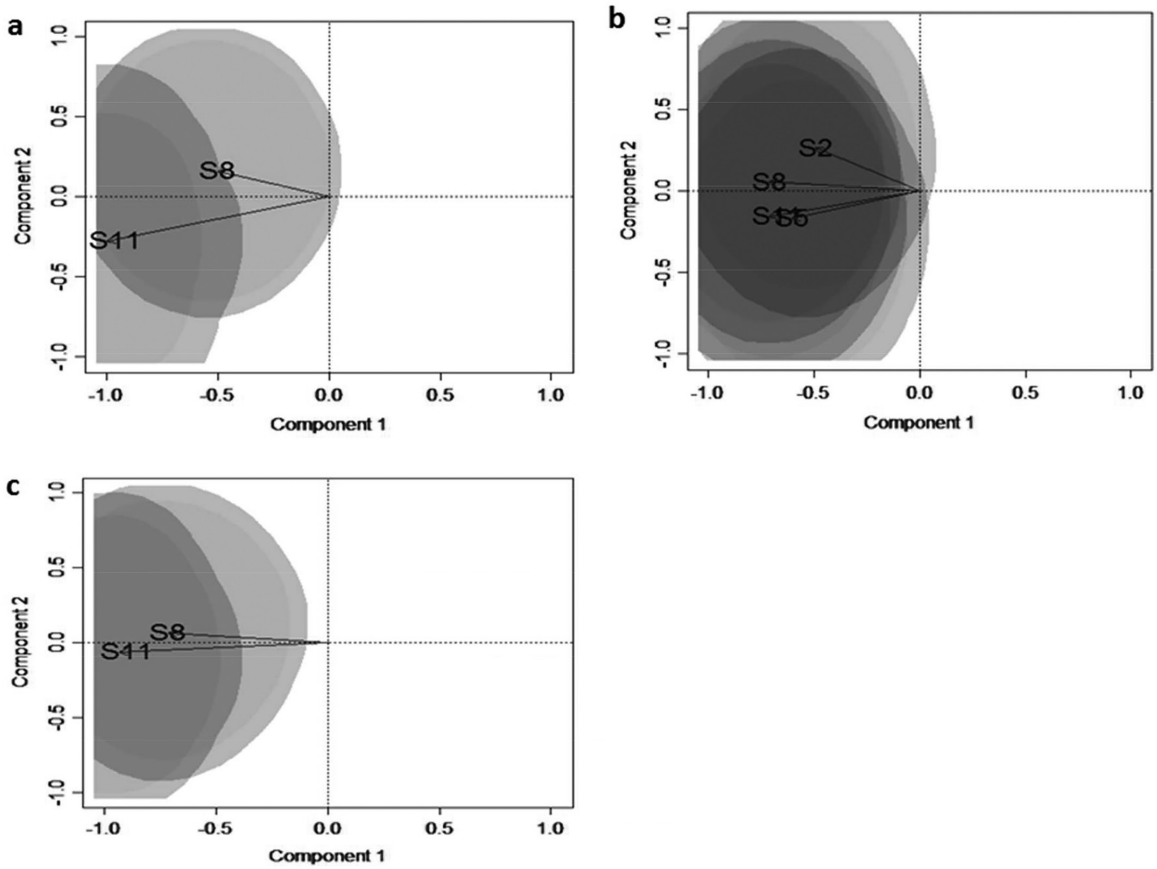

Fig. 3. Experimental data of 20 genotypes in 12 sites; plot of the column (sites) scores $VD^{1/2}$ and the bivariate 95% (gray external contour) and 90% (gray internal contour) highest posterior density (HPD) regions (only the sites that do not include the null point (0,0) at the 0.90 HPD probability level are depicted) for: (a) superpopulation, Sites S8 and S11; (b) analysis of the second evaluation period (Year 2), Sites S2, S5, S8, and S11; (c) analysis of the third evaluation period (Year 3), Sites S8 and S11.

in the center of Supplemental Fig. S2b for the analysis of Year 2 but not in Fig. 3b–c for the analysis of Years 2 and 3, respectively; therefore, under the standard analysis of Year 1, S11 did not cause $G \times E$, but under the proposed hierarchical Bayesian model, it did cause $G \times E$, as shown in Fig. 3b. Thus, the influence of all the years is simultaneously considered in the hierarchical Bayesian model.

Implications of the Hierarchical Bayesian Model for Plant Breeding Trials

The hierarchical Bayesian model is a parsimonious model, with the first component explaining nearly 95% of the variability as a result of genotypes and to $G \times E$ at both levels (yearly and superpopulation levels). No direct comparison between these results and those obtained using Bayes AMMI (Perez-Elizalde et al., 2012) can be made because the Bayes AMMI is only concerned with the $G \times E$ term in each period, whereas the hierarchical Bayesian version of model [2] estimates the combination of the main effects of genotype plus the $G \times E$ effects. Further, hierarchical Bayesian model [2] allows detecting the most important $G \times E$ in agriculture and plant breeding: crossover $G \times E$ (Crossa et al., 2004). This is important when drawing conclusions using biplots.

Breeders can benefit from the hierarchical Bayesian model because it allows them to consider information from all sites and all evaluation periods (years) when selecting the most productive and stable germplasm across the different years. Standard individual year analyses can be outperformed by superpopulation estimated parameters, which allow borrowing of information between several evaluation periods. This can be exploited by breeders when selecting germplasm that in some years may be specifically adapted to environmental conditions that are not repeated in another year. The standard variance component model allows assessing the components of variance because of main effects and two- and three-way interactions but does not provide insights into three-way genotype × site × year interaction. The hierarchical Bayesian model allows extracting information on three-way interaction because between-years correlations were implicitly considered in the hierarchical model. This offers breeders the advantage that they do not need to force combining factors, such as sites, management, and years, into one factor (i.e., environment) and can thus perform the standard two-way linear–bilinear fixed effect model.

Based on these results, we thought it was appropriate to make inferences for interaction parameters based only

Table 3. Experimental data. For the second evaluation period (year 2), the posterior summary (mean), standard deviation (SD), quartile ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.90 highest posterior density (HPD), and 0.95 HPD intervals were computed using 10,000 approximately independent samples simulated from the joint posterior distribution of the residual variance ($\sigma = \tau^{-1/2}$), all the singular values ($\lambda_1,\ldots,\lambda_{11}$), and the left and right singular vector elements of genotypes (u_{ik}) and sites (v_{jk}), respectively, whose 0.90 HPD and 0.95 HPD intervals do not contain the null value (0,0). Grain yield data were measured in kilograms per hectare.

						0.90 HPD interval		0.95 HPD interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	Lower	Upper	Lower	Upper
σ	463.7	12.9	455.5	463.7	473.0	444.6	486.8	440.2	490.9
λ_1	2354.6	502.1	2094.9	2413.8	2689.0	1581.8	3129.0	1345.0	3282.1
λ_2	969.7	494.3	600.5	960.5	1323.0	99.3	1711.8	0.2	1802.0
λ_3	470.0	334.5	193.0	420.4	689.0	0.0	941.6	0.0	1088.3
λ_4	228.0	216.0	56.3	163.6	342.1	0.0	541.3	0.0	667.7
λ_5	110.9	131.8	16.9	61.3	157.2	0.0	294.3	0.0	384.5
λ_6	53.9	78.3	5.1	22.1	69.1	0.0	152.2	0.0	216.7
λ_7	26.4	45.9	1.6	8.2	30.3	0.0	74.5	0.0	118.4
λ_8	13.0	26.4	0.5	3.0	12.8	0.0	36.0	0.0	61.3
λ_9	6.5	15.4	0.2	1.1	5.4	0.0	17.5	0.0	31.8
λ_{10}	3.2	8.8	0.1	0.4	2.2	0.0	8.0	0.0	15.7
λ_{11}	1.6	5.0	0.0	0.1	0.9	0.0	3.7	0.0	7.8
$u_{1,1}$	−0.549	0.152	−0.649	−0.580	−0.489	−0.761	−0.340	−0.784	−0.234
$u_{2,1}$	−0.281	0.140	−0.374	−0.296	−0.206	−0.499	−0.058	−0.542	0.004
$u_{5,1}$	0.300	0.138	0.224	0.312	0.391	0.085	0.519	0.020	0.560
$u_{17,1}$	0.227	0.122	0.156	0.235	0.309	0.032	0.423	−0.017	0.466
$v_{2,1}$	−0.320	0.137	−0.410	−0.330	−0.244	−0.535	−0.114	−0.575	−0.055
$v_{3,1}$	0.245	0.151	0.161	0.257	0.344	0.009	0.478	−0.046	0.534
$v_{4,1}$	0.241	0.132	0.166	0.248	0.327	0.039	0.459	−0.023	0.493
$v_{5,1}$	−0.344	0.140	−0.436	−0.357	−0.269	−0.559	−0.134	−0.609	−0.078
$v_{8,1}$	−0.412	0.127	−0.492	−0.423	−0.347	−0.609	−0.235	−0.646	−0.173
$v_{11,1}$	−0.404	0.138	−0.491	−0.416	−0.334	−0.626	−0.212	−0.663	−0.143

Table 4. Experimental data. For the third evaluation period (Year 3), the posterior summary (mean), standard deviation (SD), quartile ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.90 highest posterior density (HPD), and 0.95 HPD intervals were computed using 10,000 approximately independent samples simulated from the joint posterior distribution of the residual variance ($\sigma = \tau^{-1/2}$), all the singular values ($\lambda_1,\ldots,\lambda_{11}$) and the left and right singular vector elements of genotypes (u_{ik}) and sites (v_{jk}), respectively, whose 0.90 HPD and 0.95 HPD intervals do not contain the null value (0, 0). Grain yield data were measured in kilograms per hectare.

						0.90 HPD interval		0.95 HPD interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	Lower	Upper	Lower	Upper
σ	483.9	12.9	476.2	484.5	493.3	464.2	506.4	460.8	511.6
λ_1	2392.9	501.8	2146.4	2455.9	2724.3	1665.5	3203.2	1343.5	3294.3
λ_2	906.4	480.4	539.5	887.0	1241.3	18.8	1570.2	0.8	1731.8
λ_3	421.1	316.2	162.7	363.2	617.2	0.0	869.6	0.0	1023.5
λ_4	197.6	193.8	47.1	136.9	291.6	0.0	472.4	0.0	596.5
λ_5	93.8	114.3	13.6	50.5	131.2	0.0	247.5	0.0	334.8
λ_6	45.3	67.1	4.2	18.2	57.7	0.0	126.8	0.0	184.0
λ_7	22.2	39.0	1.3	6.7	25.0	0.0	62.5	0.0	98.0
λ_8	11.0	22.9	0.4	2.4	10.6	0.0	30.3	0.0	50.8
λ_9	5.5	13.7	0.1	0.9	4.5	0.0	14.6	0.0	26.2
λ_{10}	2.7	7.9	0.0	0.3	1.8	0.0	6.7	0.0	13.0
λ_{11}	1.3	4.6	0.0	0.1	0.7	0.0	3.1	0.0	6.4
$u_{1,1}$	−0.420	0.137	−0.510	−0.441	−0.356	−0.622	−0.218	−0.660	−0.141
$u_{2,1}$	−0.365	0.156	−0.468	−0.387	−0.289	−0.600	−0.126	−0.642	−0.046
$u_{15,1}$	0.224	0.125	0.153	0.234	0.307	0.031	0.424	−0.040	0.451
$u_{17,1}$	0.251	0.127	0.179	0.261	0.335	0.049	0.452	−0.010	0.488
$v_{2,1}$	−0.253	0.133	−0.337	−0.258	−0.177	−0.464	−0.051	−0.516	−0.004
$v_{3,1}$	0.222	0.138	0.143	0.232	0.311	0.009	0.447	−0.057	0.483
$v_{6,1}$	0.288	0.126	0.219	0.297	0.370	0.102	0.491	0.040	0.520
$v_{8,1}$	−0.409	0.124	−0.488	−0.418	−0.344	−0.601	−0.236	−0.632	−0.175
$v_{11,1}$	−0.521	0.139	−0.608	−0.539	−0.463	−0.717	−0.343	−0.751	−0.257

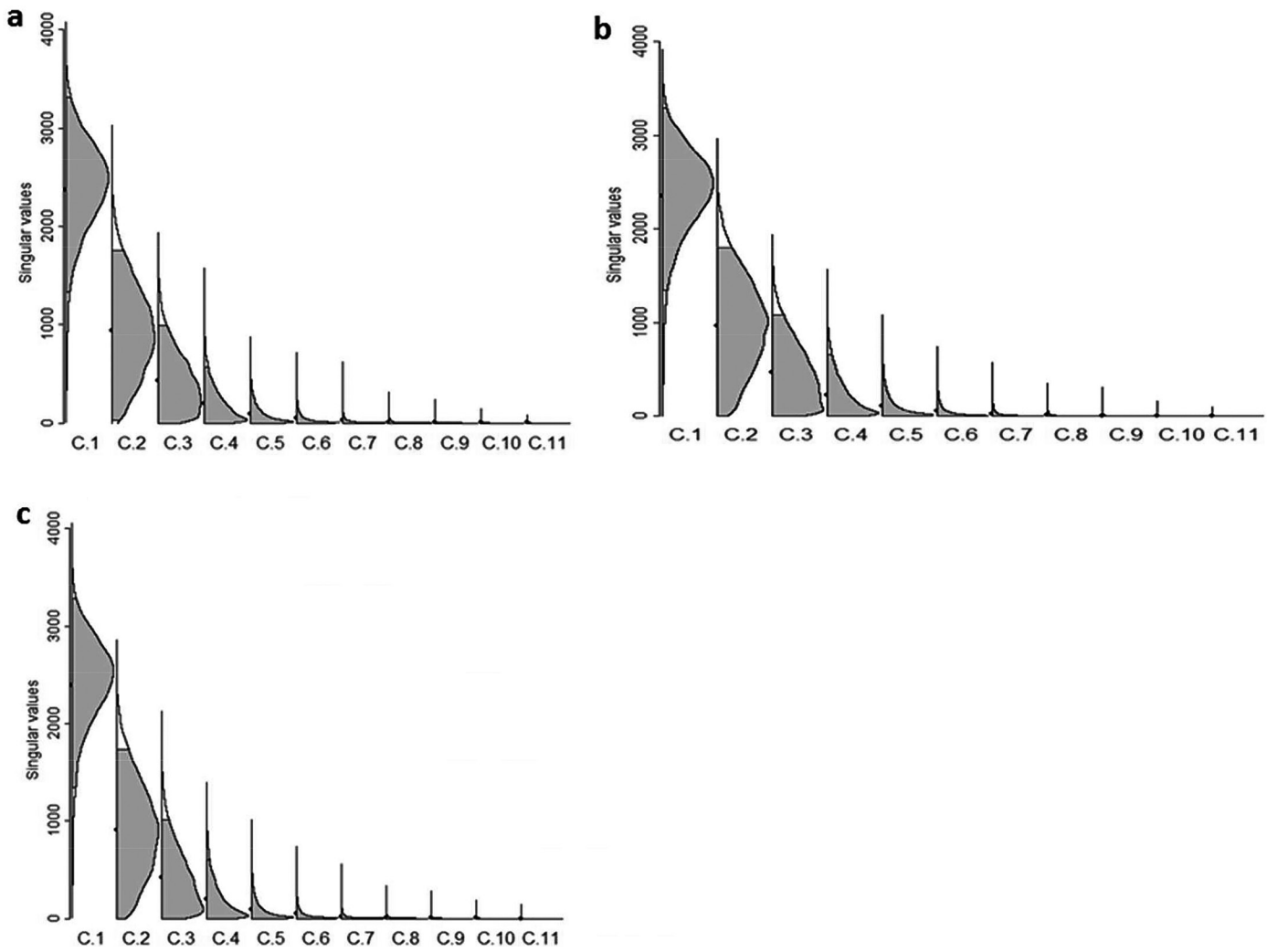

Fig. 4. Experimental data of 20 genotypes in 12 sites; posterior densities and 0.95 highest posterior density (HPD) regions of the singular values $\lambda_1,\ldots,\lambda_{11}$ (C.1–C.11) for: (a) superpopulation; (b) analysis of the second evaluation period (Year 2); (c) analysis of the third evaluation period (Year 3).

on the first two components and not to test the significance of the bilinear terms using the Bayes factor as performed by Perez-Elizalde et al. (2012) for the Bayesian AMMI. For hierarchical Bayesian analysis, estimating the Bayes factor with the MCMC sample would be a very demanding computational procedure. Also, the number of models increases not only because the number of components in the superpopulation is tested, but also because this test has to be performed for each evaluation period (year). Based on the fact that two components were used to construct the biplot and that, in this example, they explained a sizeable proportion of the variance, we decided to consider two as the number of components needed to make inferences on $G \times E$ parameters.

Evaluating newly developed crop varieties under different environmental conditions in different years is becoming increasingly important given the drastic seasonal changes in climate occurring in various regions of the world. The current global climatic situation is creating an urgent need to breed stable crop varieties that can cope with sudden climatic changes (spatial and temporal) without drastically reducing their overall productivity. The hierarchical Bayesian method can be employed to assess the complex differential responses of genotypes to temporal as well as spatial variability and could contribute to accelerating the breeding process.

CONCLUSIONS

In this research, a Bayesian inference was proposed for analyzing linear–bilinear models by considering the multivariate von Mises-Fisher distribution as prior distribution for parameters associated with $G \times E$. Simulation and real data show that the proposed hierarchical Bayesian model is an efficient tool for incorporating information from several evaluation periods when analyzing multisite trials, producing accurate results not only for individual periods, but also for superpopulation parameters.

Results presented here show that hierarchical Bayesian analysis is appropriate for dealing with differences between evaluation periods because all available information (prior

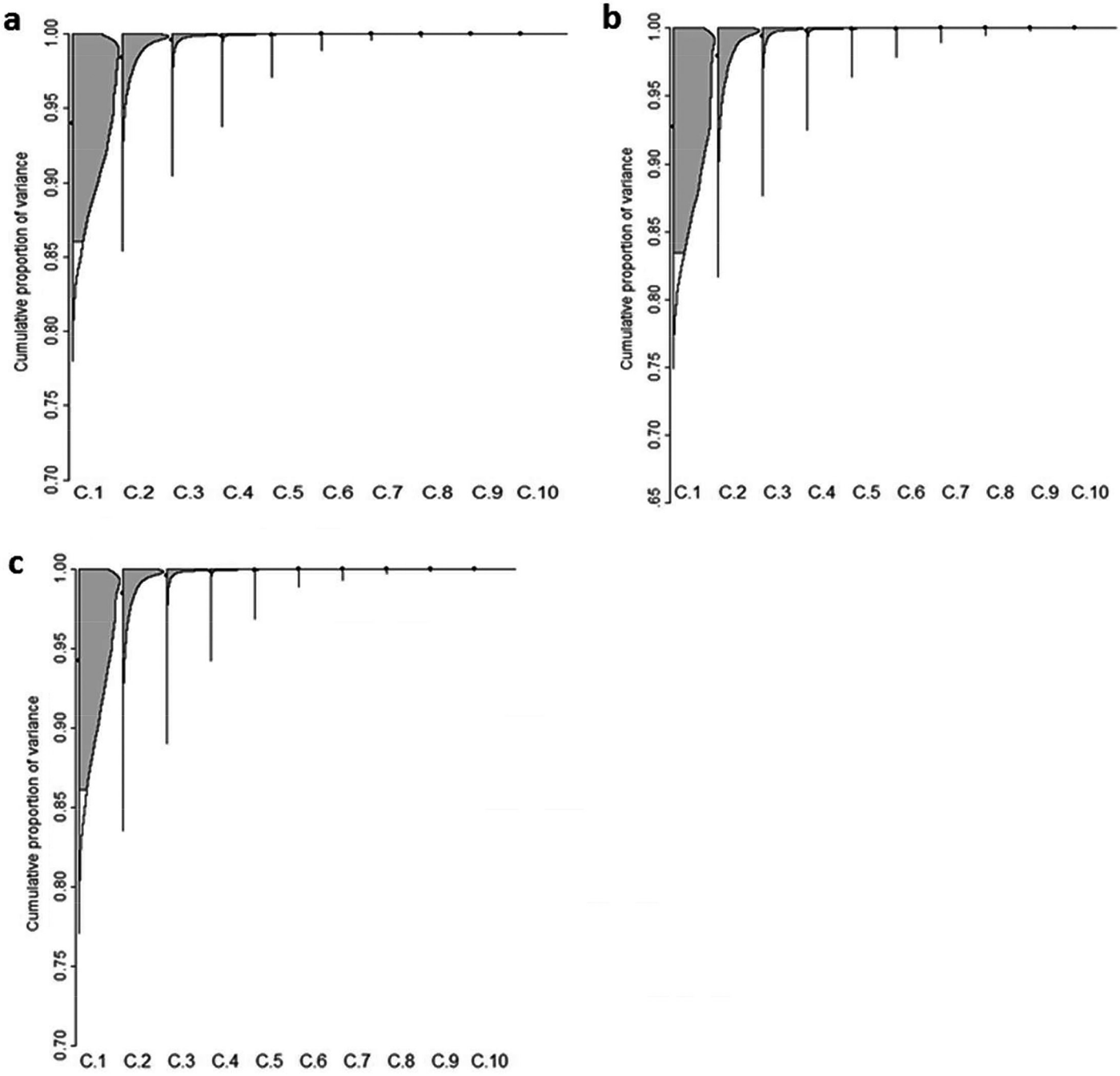

Fig. 5. Experimental data of 20 genotypes in 12 sites; posterior densities and 0.95 highest posterior density (HPD) regions of the cumulative proportion of variances $\phi_t = \frac{\sum_{k=1}^{t} \lambda_k^2}{\sum_{k=1}^{\min(r,c)-1} \lambda_k^2}$, t = 1,..,min(r,c) – 2 for: (a) superpopulation; (b) analysis of the second evaluation period (Year 2); (c) analysis of the third evaluation period (Year 3).

information and information from each evaluation period) potentially could be included in the analysis simultaneously. Because the superpopulation model naturally incorporates all available information from periods (years), the corresponding posterior distribution has a pooled precision parameter that improves its accuracy as the number of periods increases. Also, the estimation of unknown superpopulation parameters can be viewed as the realization of the joint performance of all individual periods.

Variation in the response over years has been much discussed, and some inconsistencies in the results from one period (year) to another have frequently been found. As was shown, including several data sources helps to overcome these inconsistencies in the response as a result of natural variations (usually considered as a year effect). The robustness of the hierarchical Bayesian model arises mainly from the borrowing of information between periods through superpopulation parameters and potentially with the inclusion of prior information. However,

further studies and research are needed to examine how the approach proposed by Josse et al. (2014) for dealing with overparameterization and model constraints could be implemented with a hierarchical Bayesian inference for multiyear and multisite plant breeding trials. Also, further research is required to extend the hierarchical Bayesian model presented in this study to imbalanced cases where some genotypes are not present in every site–year combination or some sites are not included in some years.

APPENDIX A
Full Conditional Posterior Distributions

$$\pi\left(\theta_m^* \mid \tau_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \mathbf{Y}_m\right) = \mathrm{N}_s\left(\theta_m^* \mid \left(\Sigma_{am}^{*-1} + \Sigma_m^{*-1}\right)^{-1}\left(\Sigma_{am}^{-1}\hat{\theta}_m^* + \Sigma_m^{*-1}\theta^*\right), \left(\Sigma_{am}^{*-1} + \Sigma_m^{*-1}\right)^{-1}\right)$$

where $\hat{\theta}_m^* = \left(\frac{n_m 1_{\boldsymbol{g}}' \mathbf{Y}_m 1_s}{n_m gs}, \frac{n_m \mathbf{K}_s' \mathbf{Y}_m' 1_g}{g}\right)$ and are the estimations of the linear terms and the singular covariance matrix for the current period under study.

$$\pi\left(\mathbf{D}_m \mid \tau_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{Y}_m, \mathbf{D}\right) = \prod_{k=1}^{t}\left\{1 - \phi\left(\sqrt{\left(\tau_m\left(n_m + n\right)\right)^{-1}}\left(\lambda_{(k+1)_m} - \frac{n_m l_{k_m} + n l_k}{n_m + n}\right)\right)\right\}^{-1} \times$$

$$\mathrm{N}\left(\lambda_{k_m} \mid \frac{n_m l_{k_m} + n l_k}{n_m + n}, \left(\tau_m\left(n_m + n\right)\right)^{-1}\right)$$

$$\lambda_{1_m} > \lambda_{2_m} > \ldots > \lambda_{t_m} > \lambda_{(t+1)_m} = 0$$

$$\pi\left(\mathbf{U}_m \mid \tau_m, \mathbf{V}_m, \mathbf{D}_m, \mathbf{Y}_m, \mathbf{M}_{0m}, \mathbf{V}, \mathbf{D}\right) \propto \mathrm{etr}\left(\tau_m\left[n_m \mathbf{Y}_m \mathbf{V}_m \mathbf{D}_m + n\mathbf{M}_{0m}\mathbf{V}\mathbf{D}\right]\mathbf{U}'_m\right)$$

$$\pi\left(\mathbf{V}_m \mid \tau_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{Y}_m, \mathbf{M}_{0m}, \mathbf{U}, \mathbf{D}\right) \propto \mathrm{etr}\left(\tau_m\left[n_m \mathbf{Y}'_m \mathbf{U}_m \mathbf{D}_m + n\mathbf{M}'_{0m}\mathbf{U}\mathbf{D}\right]\mathbf{V}'_m\right)$$

$$\pi\left(\tau_m \mid \mathrm{a}_m, \mathrm{b}_m, \mathbf{E}_m\right) = \mathrm{Gamma}\left(\tau_m \mid \frac{n_m rc}{2} + \mathrm{a}_m, \frac{n_m}{2}\mathrm{tr}\left(\mathbf{E}_m\mathbf{E}'_m\right) + \frac{\left(n_m - 1\right)\mathrm{tr}\left(\mathbf{S}_m\mathbf{S}'_m\right)}{2} + \mathrm{b}_m\right)$$

$$\pi\left(\theta^* \mid \tau_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \sum_{m=1}^{h}\Sigma_m^{*-1}, \mathbf{Y}_m\right) = \mathrm{N}\left(\theta^* \mid \left(\sum_{m=1}^{h}\Sigma_m^{*-1} + \Sigma_p^{*-1}\right)^{-1}\left(\sum_{m=1}^{h}\Sigma_m^{*-1}\theta_m^* + \Sigma_p^{*-1}\theta_0^*\right), \left(\sum_{m=1}^{h}\Sigma_m^{*-1} + \Sigma_p^{*-1}\right)^{-1}\right)$$

$$\pi\left(\mathbf{D} \mid \tau_m, \mathbf{D}_m\right) =$$

$$\prod_{k=1}^{t}\left\{1 - \phi\left(\sqrt{\left(n\sum_{m=1}^{h}\tau_m + n_0\tau_0\right)^{-1}}\left(\lambda_{k+1} - \frac{n\sum_{m=1}^{h}\tau_m l_{k_m} + n_0\tau_0 l_k^0}{n\sum_{m=1}^{h}\tau_m + n_0\tau_0}\right)\right)\right\}^{-1} \times$$

$$\mathrm{N}\left(\lambda_k \mid \frac{n\sum_{m=1}^{h}\tau_m l_{k_m} + n_0\tau_0 l_k^0}{n\sum_{m=1}^{h}\tau_m + n_0\tau_0}, \left(n\sum_{m=1}^{h}\tau_m + n_0\tau_0\right)^{-1}\right)$$

$$\lambda_1 > \lambda_2 > \ldots > \lambda_t > \lambda_{t+1} = 0$$

$$\pi\left(\mathbf{U} \mid \tau_m, \mathbf{V}_m, \mathbf{M}_{0m}, \mathbf{D}\right) \propto \text{etr}\left(\left[n\sum_{m=1}^{h}\tau_m \mathbf{M}_{0m}\mathbf{V}\mathbf{D} + n_0\tau_0\mathbf{M}_0\mathbf{V}_0\mathbf{D}_0 \right]\mathbf{U}'\right)$$

$$\pi\left(\mathbf{V} \mid \tau_m, \mathbf{U}_m, \mathbf{M}_{0m}, \mathbf{D}\right) \propto \text{etr}\left(\left[n\sum_{m=1}^{h}\tau_m \mathbf{M}'_{0m}\mathbf{U}\mathbf{D} + n_0\tau_0\mathbf{M}'_0\mathbf{U}_0\mathbf{D}_0 \right]\mathbf{V}'\right)$$

APPENDIX B

Table B1. Experimental data. Analysis of the superpopulation, posterior summary (mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$) and 0.90 highest posterior density (HPD), and 0.95 HPD intervals of the 10,000 approximately independent samples simulated from the joint posterior distribution of the 13 linear effects measured in kilograms per hectare.

						0.90 HPD interval		0.95 HPD interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	Lower	Upper	Lower	Upper
μ	5004.4	21.6	4989.8	5004.3	5019.1	4970.0	5040.8	4963.1	5047.1
β_1	2285.0	72.3	2236.4	2284.8	2333.7	2167.2	2405.1	2143.1	2424.5
β_2	2366.9	72.8	2317.8	2366.0	2415.4	2246.1	2486.0	2221.3	2507.4
β_3	2273.5	71.5	2224.3	2273.1	2322.4	2157.5	2390.8	2137.5	2416.5
β_4	−878.8	72.1	−927.6	−878.8	−830.5	−1000.0	−763.2	−1015.8	−733.8
β_5	−501.8	71.9	−550.0	−502.2	−453.0	−619.1	−382.8	−642.8	−361.0
β_6	−776.0	72.4	−824.9	−776.7	−726.2	−892.2	−655.6	−912.9	−631.2
β_7	−812.9	75.3	−863.7	−813.0	−761.7	−939.7	−692.8	−961.1	−665.6
β_8	−153.5	71.7	−201.8	−153.6	−105.8	−268.3	−33.7	−289.9	−8.8
β_9	−173.7	71.7	−222.1	−173.3	−125.4	−293.1	−59.0	−314.5	−33.8
β_{10}	−1161.3	72.1	−1209.4	−1161.5	−1112.9	−1283.4	−1046.0	−1300.3	−1016.6
β_{11}	−978.9	71.8	−1027.8	−978.1	−930.1	−1097.2	−861.7	−1119.8	−838.7
β_{12}	−1488.5	73.8	−1538.1	−1488.9	−1439.1	−1609.7	−1367.4	−1629.2	−1339.5

Table B2. Experimental data. Analysis of the second evaluation period (Year 2), posterior summary (mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$) and 0.90 highest posterior density (HPD), and 0.95 HPD intervals of the 10,000 approximately independent samples simulated from the joint posterior distribution of 13 linear effects of grain yield measured in kilograms per hectare.

						0.90 HPD interval		0.95 HPD Interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	Lower	Upper	Lower	Upper
μ	4981.4	18.5	4968.9	4981.3	4993.8	4950.6	5011.5	4944.7	5016.8
β_1	2423.4	60.9	2382.0	2423.6	2464.5	2322.7	2522.6	2304.2	2542.2
β_2	2576.7	61.4	2534.9	2576.9	2618.6	2473.4	2674.4	2461.0	2699.1
β_3	2302.3	61.6	2261.1	2302.5	2343.8	2202.8	2405.4	2186.3	2428.2
β_4	−956.7	61.4	−998.6	−957.0	−915.3	−1054.4	−853.7	−1078.2	−837.9
β_5	−464.0	61.2	−505.1	−464.1	−422.9	−562.5	−361.2	−584.0	−343.5
β_6	−899.6	61.7	−941.4	−899.7	−858.2	−1000.9	−798.0	−1020.8	−779.2
β_7	−480.1	62.2	−522.3	−480.1	−438.0	−580.2	−375.7	−601.4	−359.4
β_8	−100.9	61.0	−142.3	−101.2	−59.7	−202.2	−2.7	−221.1	18.6
β_9	−253.4	61.6	−294.7	−253.5	−212.1	−356.5	−154.9	−371.8	−129.7
β_{10}	−1311.6	61.5	−1353.0	−1311.6	−1270.2	−1414.6	−1213.2	−1433.2	−1190.8
β_{11}	−1080.4	61.3	−1121.7	−1080.1	−1039.3	−1183.4	−982.6	−1201.0	−961.2
β_{12}	−1755.7	62.1	−1797.8	−1755.3	−1714.3	−1861.1	−1656.7	−1880.2	−1636.4

Table B3. Experimental data. Analysis of the third evaluation period (Year 2), posterior summary (mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$) and 0.90 highest posterior density (HPD), intervals of the 10,000 approximately independent samples simulated from the joint posterior distribution of 13 linear effects of grain yield measured in kilograms per hectare.

						0.90 HPD interval		0.95 HPD interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	Lower	Upper	Lower	Upper
μ	5029.6	19.0	5016.7	5029.5	5042.4	4998.5	5060.7	4992.5	5067.0
β_1	2133.9	63.2	2091.3	2133.7	2176.4	2031.0	2238.9	2010.7	2257.1
β_2	2138.3	63.0	2096.4	2138.1	2180.4	2036.7	2243.0	2013.4	2260.9
β_3	2241.9	63.2	2199.5	2242.2	2284.6	2137.3	2344.3	2116.5	2363.3
β_4	–793.3	62.9	–835.7	–793.2	–750.9	–897.0	–691.3	–916.3	–669.0
β_5	–543.2	63.4	–586.1	–543.4	–500.3	–649.8	–441.3	–664.6	–417.1
β_6	–641.6	63.5	–684.1	–641.5	–598.6	–746.3	–537.3	–769.7	–520.4
β_7	–1176.3	64.0	–1219.5	–1176.0	–1133.6	–1280.3	–1069.9	–1306.3	–1055.7
β_8	–211.7	62.7	–254.0	–212.1	–169.6	–314.0	–108.7	–332.1	–85.8
β_9	–86.0	63.2	–128.1	–86.4	–43.4	–191.2	15.7	–213.4	35.1
β_{10}	–997.5	63.0	–1040.0	–997.6	–955.3	–1101.0	–894.1	–1125.8	–878.7
β_{11}	–867.4	63.0	–910.1	–867.4	–825.0	–968.3	–761.5	–990.8	–744.5
β_{12}	–1197.3	63.7	–1240.5	–1197.5	–1153.6	–1299.8	–1091.6	–1324.0	–1076.1

Acknowledgments

The first author would like to acknowledge the financial support provided by CONACYT (Mexico's Consejo Nacional de Ciencia y Tecnología). All four authors would like to thank the national cooperators who carried out the multienvironment, multiyear maize trials included in this study.

References

Cornelius, P.L., and J. Crossa. 1999. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment cultivar trials. Crop Sci. 39:998–1009. doi:10.2135/cropsci1999.0011183X003900040007x

Cornelius, P.L., J. Crossa, M.S. Seyedsadr, G. Liu, and K. Viele K. 2001. Contributions to multiplicative model analysis of genotype–environment data. Proc. of the Annu. Meeting of the Am. Statistical Assoc., Atlanta, GA. 5–9 Aug. 2001.

Cornelius, P.L., and M.S. Seyedsadr. 1997. Estimation of general linear–bilinear models for two-way tables. J. Stat. Comput. Simul. 58:287–322. doi:10.1080/00949659708811837

Cotes, J.M., J. Crossa, A. Sanches, and P.L. Cornelius. 2006. A Bayesian approach for assessing the stability of genotype. Crop Sci. 46:2654–2665. doi:10.2135/cropsci2006.04.0227

Crossa, J., and P.L. Cornelius. 1997. Site regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. Crop Sci. 37:406–415. doi:10.2135/cropsci1997.0011183X003700020017x

Crossa, J., S. Perez-Elizalde, D. Jarquin, M. Cotes, K. Viele, G. Liu, and P.L. Cornelius. 2011. Bayesian estimation of the additive main effects and multiplicative interaction model. Crop Sci. 51:1458–1469. doi:10.2135/cropsci2010.06.0343

Crossa, J., R.C. Yan, and P.L. Cornelius. 2004. Studying crossover genotype × environment interaction using linear–bilinear models and mixed models. J. Agric. Biol. Environ. Stat. 9:362–380. doi:10.1198/108571104X4423

da Silva, C.P., L.A. de Oliveira, J.J. Nuvunga, A.K.A. Pamplona, and M. Balestre. 2015. A Bayesian Shrinkage Approach for AMMI Models. PLoS ONE 10:E0131414. doi:10.1371/journal.pone.0131414

Edwards, J.W., and J.L. Jannink. 2006. Bayesian modeling of heterogeneous error and genotype × environment interaction variances. Crop Sci. 46:820–833. doi:10.2135/cropsci2005.0164

Foucteau, V., and J.B. Denis. 2000. Statistical analysis of successive series of experiments in plant breeding: A Bayesian approach. In: A. Gallais, C. Dillmann, and I. Goldringer, editors, Eucarpia, quantitative genetics and breeding methods: The way ahead. INRA Editions, Versailles, France. p. 49–56.

Gabriel, K.R. 1978. Least squares approximation of matrices by additive and multiplicative models. J. R. Stat. Soc. Series B Stat. Methodol. 40:186–196.

Gelman, A. 2004. Bayesian data analysis. Chapman and Hall/CRC, Boca Raton, FL.

Gelman, A., and D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7:457–472. doi:10.1214/ss/1177011136

Hoff, P.D. 2009. Simulation of the matrix bingham–von mises–fisher distribution, with applications to multivariate and relational data. J. Comput. Graph. Stat. 18:438–456. doi:10.1198/jcgs.2009.07177

Josse, J., F. van Eeuwijk, H.P. Piepho, and J.D. Denis. 2014. Another look at Bayesian analysis of AMMI models for genotype–environment data. J. Agric. Biol. Environ. Stat. 19:240–257 doi:10.1007/s13253-014-0168-z.

Liu, G. 2001. Bayesian computations for general linear–bilinear models. Ph.D. diss., Dep. of Statistics, Univ. of Kentucky, Lexington.

Paderewski, P., H.G. Gauch, Jr., W. Mądry, and. E. Gacek. 2015. AMMI analysis of four-way genotype × location × management × year data from a wheat trial in Poland. Crop Sci. 55:1–8 doi:10.2135/cropsci2015.03.0152

Perez-Elizalde, S., D. Jarquin, and J. Crossa. 2012. A general Bayesian estimation method of linear–bilinear models applied to plant breeding trials with genotype × environment interaction. J. Agric. Biol. Environ. Stat. 17:15–37. doi:10.1007/s13253-011-0063-9

Raftery, A.E., and S.M. Lewis. 1995. The number of iterations, convergence diagnostics and generic metropolis algorithms. In: W.R. Gilks and D.J. Spiegelhalter, editors, Practical Markov Chain Monte Carlo. Chapman and Hall, London, p. 115–130.

Shukla, G.K. 1972. Some statistical aspects of partitioning genotype–environment components of variability. Heredity 29:237–245. doi:10.1038/hdy.1972.87

Theobald, C., M. Talbot, and F. Nabugoomu. 2002. A Bayesian approach to regional and local-area prediction from crop variety trials. J. Agric. Biol. Environ. Stat. 7:403–419. doi:10.1198/108571102230

Varela, M., J. Crossa, J. Rane, A. Joshi, and R. Threthowan. 2006. Analysis of a three-way interaction including multi-attributes. Aust. J. Agric. Res. 57:1185–1193. doi:10.1071/AR06081

Viele, K., and C. Srinivasan. 2000. Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback–Leibler information. J. Stat. Plan. Inference 84:201–219. doi:10.1016/S0378-3758(99)00151-2