COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

Análisis Bayesiano de Modelos Lineales - Bilineales

JUAN DIEGO HERNÁNDEZ JARQUÍN

T E S I S

PRESENTADA COMO REQUISITO PARCIAL

PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO

FEBRERO 2012

La presente tesis titulada: **"Análisis Bayesiano de Modelos Lineales – Bilineales"** realizada por el alumno: **Juan Diego Hernández Jarquín,** bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTOR EN CIENCIAS
SOCIOECONOMÍA ESTADÍSTICA E INFORMÁTICA

ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO _____
DR. SERGIO PEREZ ELIZALDE

ASESOR _____
DR. FELIX V. GONZALEZ COSSIO

ASESOR _____
DR. JUAN ANDRES BURGUEÑO FERREIRA

ASESOR _____
DR. JOSE LUIS CROSSA HIRTART

ASESOR _____
DR. GUSTAVO RAMIREZ VALVERDE

Montecillo, Texcoco, Estado de México, Febrero 2012

Análisis Bayesiano de Modelos Lineales - Bilineales

Juan Diego Hernández Jarquín, Doctor.

Colegio de Postgraduados, 2012.

El análisis de tablas de doble entrada es una herramienta estadística que se presenta en diversos campos de investigación; por ejemplo, en fitomejoramiento uno de los principales objetivos es evaluar la adaptabilidad y estabilidad genotípica en la selección de los padres para el siguiente ciclo de mejoramiento. Generalmente, este proceso se ve afectado por la presencia de la interacción Genotipo x Ambiente (GE). Bajo el enfoque clásico, para el estudio de la interacción se consideran modelos parsimoniosos como el AMMI ó el SREG y se obtienen estimaciones puntuales mediante Mínimos Cuadrados Ordinarios (MCO) por lo que no es trivial la construcción de intervalos de confianza y el diseño de pruebas de hipótesis. En este trabajo se propone una modelación bayesiana de los modelos lineales-bilineales que ofrece la ventaja de incorporar información a priori, con este enfoque se obtienen estimaciones puntuales encogidas de los eigenvalores. Por otro lado, una vez que se obtiene la distribución a posteriori es posible el cálculo de regiones bivariadas de alta probabilidad a posteriori (HPD) y de regiones de credibilidad para los parámetros scores; también es factible el diseño de pruebas de hipótesis bayesianas, a través de los factores Bayes, sobre el número de términos bilineales que debe contener el modelo. Para las matrices singulares derivadas de la descomposición en valores singulares de la matriz de interacción se propone como distribución a priori la distribución von Mises Fisher vectorial. La organización de este trabajo se divide en tres Capítulos. En el Capitulo 1 se propone el modelo AMMI bayesiano haciendo uso de distribuciones a priori no informativas; en el Capítulo 2 se plantea una formulación matricial del modelo AMMI bayesiano que ofrece la ventaja de incorporar información a priori sobre la interacción por medio de una matriz de medias a priori; el Capitulo 3 desarrolla un modelo jerárquico bayesiano cuya principal ventaja es el incorporar información de una serie de experimentos.

Palabras clave: Distribución von Mises-Fisher; Inferencia Bayesiana; Fitomejoramiento; Tablas de doble entrada con interacción; Términos Bilineales de interacción.

Bayesian Analysis of Linear – Bilinear Models

Juan Diego Hernández Jarquín, Doctor.

Colegio de Postgraduados, 2012.

The two-way table analysis is a useful tool that arises in many fields of research; for example in plant breeding the main purpose is to asses genotypic adaptability and stability that will allow make an accurate selection of parents for the next breeding cycle. The presence of Genotype x Environmental Interaction (GE) complicates this process. Generally, the study of this interaction has been conducted using the least square method in parsimonious models, as the AMMI model and the SREG model, yielding punctual estimates. For this reason, is not trivial the construction of confidence intervals neither the design of hypothesis testing. This research proposed a bayesian modelation of the linear-bilinear models which offers advantages as incorporate prior information; this approach yields shrinkage estimates of the eigenvalues. By the other hand, the posterior distribution allows obtain bivariate highest posterior density (HPD) regions and credible intervals for the score parameters, design of bayesian hypothesis testing for determinate the number of components to be retained in the model through the use of the Bayes factor. For the singular matrices resulting from the singular value decomposition of the residual matrix the vectorial von Mises Fisher distribution is proposed as prior distribution. The structure of this document is as follows: the Chapter 1 shows the Bayesian model using noninformative priors; the Chapter 2 formulate a matrix notation of the Bayes AMMI, here is possible incorporate prior information about interaction parameters through a prior matrix of means; in the Chapter 3 a hierarchical Bayesian model is proposed, this model offers as principal advantage the incorporation of several data sources in the analysis.

Key words: Bayesian inference; Bilinear interaction terms; Plant breeding; Two-way tables with interaction; von Mises Fisher distribution.

AGRADECIMIENTOS

Al **Consejo Nacional de la Ciencia y Tecnología (CONACyT)**, por el apoyo financiero otorgado durante la realización de mis estudios Doctorales.

Al **Colegio de Postgraduados**, en especial al **Postgrado de Socioeconomía, Estadística e Informática - Estadística,** por haberme brindado la oportunidad de concluir una etapa más de mi formación profesional.

Al **Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMyT)**, por el apoyo financiero para la realización de la presente investigación.

Al **Dr. Sergio Pérez Elizalde**, por dirigir la presente investigación, dedicarle muchas horas de su valioso tiempo y además orientarme correctamente en la toma decisiones.

Al **Dr. José Luis Crossa Hiriart**, sus conocimientos y experiencia en el tema fueron determinantes para concluir la investigación.

Al **Dr. Gustavo Ramírez Valverde**, quien es uno de los pilares de mi formación académica y siempre me motivo a seguir adelante.

Al **Dr. Juan Andrés Burgueño Ferreira**, por sus sugerencias a la realización de la presente investigación.

Al **Dr. Félix V. González Cossio**, por su apoyo incondicional durante mi estancia en el COLPOS.

Al **DR. Gilberto Rendón Sánchez**, quien en ningún momento dudo en ayudarme cuando así se lo solicite.

DEDICATORIA

A Dios por darme la vida y permitirme llegar a la meta.

Con todo mi amor a mi hijo Jarquín, por ser mi gran fuerza para seguir adelante y por todos los momentos de ausencia.

A ti, que aún no tengo el placer de conocerte pero sé que día con día vas creciendo, te amo desde el primer instante.

A mi amada esposa, que a pesar de mis tantas ausencias siempre estuvo dispuesta a ayudarme en los momentos más difíciles, sin ti nunca lo habría logrado.

A mis padres, por sus consejos de superación y apoyo en la realización de mis sueños.

A mi gran familia sanguínea, consanguínea y extendida, por su apoyo moral y brindarme la oportunidad de formar parte de ella.

A todas las secretarias del Postgrado en Estadística, en especial a Emma por su apoyo incondicional durante mi estancia en el Colegio de Postgraduados.

A Raúl y Luis, entrañables amigos que siempre me tendieron la mano.

A ti, a quien tengo siempre presente en mis oraciones y nunca te olvidare^{+}.

CONTENIDO

INTRODUCCIÓN GENERAL

El análisis estadístico de las tablas de doble entrada con interacción se presenta en diferentes campos de investigación como la Medicina, Agricultura, Genética y las Ciencias Sociales; uno de los principales objetivos del estudio de estas tablas es obtener conclusiones acerca de los componentes de interacción. Por ejemplo, en la selección de variedades las instituciones dedicadas al mejoramiento genético y las compañías dedicadas a la producción de semillas establecen experimentos en múltiples ambientes (Yan et al., 2000), en este tipo de ensayos la ocurrencia de la interacción Genotipo x Ambiente (GE) es inevitable (Ceccarelli et al., 2006).

Esta interacción puede entenderse como una inconsistencia del desempeño fenotípico de los genotipos a través de los ambientes (Asfaw et al., 2009). Cuando la interacción no se presenta el promedio a través de los diferentes ambientes es un indicador adecuado. La interacción puede ser de dos tipos: (i) "non-crossover" ocurre cuando el "ranking" de los genotipos permanece constante a través de los ambientes y la interacción es significativa debido a cambios en la magnitud de la respuesta, mientras que en el tipo (ii) "crossover" sucede un cambio significativo en el "ranking" de la respuesta de un ambiente a otro (KAYA et al., 2006).

El análisis de varianza (ANOVA) es un modelo aditivo que describe eficazmente los efectos principales y sirve para determinar si él termino de interacción es una fuente de variación significativa pero no proporciona orientación alguna acerca de los genotipos ó ambientes que dan lugar a la interacción (Samonte et al., 2005). Varios modelos estadísticos se han propuesto para el análisis de la interacción, sin embargo no todos ellos han resultado ser lo suficientemente efectivos (Zobel et al., 1988).

El modelo de efectos principales aditivos e interacción multiplicativa (AMMI) y el modelo de regresión por sitios (SREG) se consideran dos poderosas herramientas para realizar análisis efectivos (Ebbon and Gauch, 2002). El modelo AMMI combina los parámetros aditivos del ANOVA tradicional con los términos multiplicativos del análisis de componentes principales (PCA). El modelo SREG es un modelo multiplicativo que absorbe el efecto principal de uno de los componentes lineales y la interacción GGE (Yan and Tinker, 2006).

El proceso de estimación en los modelos mencionados se lleva a cabo mediante el método iterativo de mínimos cuadros ordinarios (MCO) donde primero se ajusta los términos lineales ignorando los bilineales los cuales son subsecuentemente ajustados como los primeros *t* componentes de la descomposición en valores singulares (SVD) de la matriz de residuales. Estos modelos pueden llegar a ser parsimoniosos debido a que los parámetros de interacción son estimados reteniendo solo los primeros componentes (Kempton, 1984). El uso de MCO produce estimaciones puntuales por lo que no es trivial la construcción de intervalos de confianza y el diseño de pruebas de hipótesis.

En esta investigación, se propone una modelación bayesiana la cual ofrece la ventaja de poder incorporar información a priori derivada de estudios anteriores o del conocimiento de un experto en la materia. El uso de este enfoque en la modelación de la interacción GE ha sido limitado principalmente por la estructura compleja del espacio de parámetros generado por las bases ortonormales de la descomposición en valores singulares. Viele and Srinivasan (2000) fueron los primeros en proponer la estimación Bayesiana de los parámetros de GE usando Cadenas de Markov Monte Carlo (MCMC) con Metropolis-Hastings, estos autores plantean el uso de la distribución Uniforme Esférica como distribución a priori para las matrices singulares del término bilineal. La distribución Uniforme Esférica es un caso especial de la distribución von Mises Fisher (Mardia et al., 1979), Hoff (2009) mostro como obtener muestras de la distribución von Mises Fisher en la Esfera multidimensional.

La distribución marginal a posteriori es necesaria para realizar inferencia sobre los parámetros desconocidos, pero esta envuelve integración en altas dimensiones por lo cual es indispensable el uso Cadenas de Markov Monte Carlo (MCMC) a través del Muestreador de Gibbs que además ahorrar tiempo computacional vuelve estable el algoritmo.

El contenido de este trabajo comprende de 3 Capítulos: En el Capítulo 1 se introduce el modelo AMMI bayesiano mediante el uso de distribuciones a priori no informativas y un experimento de maíz de 9 genotipos evaluados en 20 ambientes sirvió como ejemplo de aplicación. Una formulación matricial del modelo AMMI bayesiano se presenta en el Capitulo 2; este modelo ofrece la ventaja de poder incorporar de manera intuitiva información disponible y mediante el uso de los factores Bayes es posible realizar pruebas de hipótesis bayesianas sobre los términos bilineales con el fin de determinar el número de componentes que debe retener el modelo, como

ejemplo se utilizo un experimento multi-anual de maíz en dos años consecutivos donde la información del primer año se utilizo como información a priori. El Capitulo 3 desarrolla un modelo jerárquico bayesiano el cual ofrece la ventaja de incorporar información de una serie de experimentos y como ejemplo de aplicación analizo un experimento de trigo en tres años consecutivos, la información del año 1 se utilizo como información a priori. En los tres capítulos se desarrolla la teoría para el modelo AMMI el cual ofrece una familia de modelos (SREG, GREG, COMM, etc.) a partir del relajar de algunas restricciones.

CAPÍTULO 1: BAYESIAN ESTIMATION OF THE ADDITIVE MAIN EFFECTS AND MULTIPLICATIVE INTERACTION (AMMI) MODEL

Diego Jarquin, José Crossa, Sergio Perez-Elizalde, José Miguel Cotes, Kert Viele, Genzhou Liu, and Paul L. Cornelius

RESUMEN

Muchas investigaciones han sido llevadas a cabo utilizando la estimación de mínimos cuadrados bajo el modelo linear-bilinear de efectos principales aditivos e interacción multiplicativa (AMMI). La principal dificultad con los modelos lineales-bilineales estándar es que la inferencia estadística de los efectos bilineales de la interacción genotipo x ambiente (GE) no puede ser incorporada fácilmente en el biplot de los dos primeros componentes. Esta investigación propone una aproximación Bayesiana para la inferencia de los parámetros del modelo AMMI usando el muestreador de Gibbs el cual ahorra el tiempo de cálculo y hace el algoritmo estable. Los datos de un experimento de maíz (*Zea mays* L.) en multiples-ambientes (METs) fue usado para ilustración. Distribuciones a priori vagas pero propias fueron utilizadas. Los resultados muestran que las cadenas de Markov de Monte Carlo (MCMC) reunieron los criterios de convergencia para todos los parámetros. Regiones bivariadas de alta densidad a posterior (HPD) para las interacciones del AMMI-bayesiano son mostradas en el Biplot de los primeros dos componentes bilineales; estas regiones ofrecen inferencia estadística de los parámetros bilineales y permiten las visualización de grupos homogéneos de ambientes y genotipos.

Abreviaciones: AIC, el criterio de información de Akaike; AMMI, efectos principales aditivos e interacción multiplicativa; BLUP, mejor predictor lineal insesgado; GE, interacción genotipo x ambiente; GGE, ambiente más interacción genotipo x ambiente; HPD, alta densidad a posterior; MET, experimentos de multiples-ambientes; MCMC, cadenas de Markov de Monte Carlo; SREG, regresión de sitios.

Capítulo publicado como artículo científico en Cropscience, 51: 1458-1469 (2011).

ABSTRACT

Much research has been conducted using least squares estimates of the linear-bilinear model Additive Main effects and Multiplicative Interaction (AMMI). The main difficulty with the standard linear-bilinear models is that statistical inference on the bilinear effects of genotype × environment interaction cannot be incorporated easily into the biplot of the first two components. This research proposes a Bayesian approach for the inference on the parameters of the AMMI model using a Gibbs sampler that saves computing time and makes the algorithm stable. Data from one maize (*Zea mays* L.) multi-environment trial (MET) was used for illustration. Vague but proper prior distributions were introduced. Results show that the various Markov chain Monte Carlo convergence criteria were met for all parameters. Bivariate Highest Posterior Density (HPD) regions for the Bayesian-AMMI interactions are shown in the biplot of the first two bilinear components; these regions offer a statistical inference on the bilinear parameters and allow visualizing homogeneous groups of environments and genotypes.

Abbreviations: AIC, Akaike´s information criterion; AMMI, additive main effects and multiplicative interaction; BLUP, best linear unbiased prediction; GE, genotype x environment interaction; GGE, genotype plus genotype x environment interaction; HPD, highest posterior density; MET, multi-environment trial; MCMC, Markov chain Monte Carlo; SREG, sites regression.

INTRODUCTION

In plant breeding, the main purposes of multi-environment trials (METs) are to: (i) study genotype × environment interaction (GE); (ii) assess genotypic adaptability and stability; (iii) establish relationships among testing environments, among genotypes, and among genotypes and environments (or sites) simultaneously; and (iv) make predictions of the genotypes' breeding value that will allow making an accurate selection of parents for the next breeding cycle. The presence of GE complicates this process and is usually expressed either as inconsistent responses of some genotypes relative to others, due to genotypic rank change, or as changes in the absolute differences between genotypes without rank change (Crossa et al., 2004).

Linear-bilinear (also called multiplicative) models for the study of two-way interactions date back to Mandel (1961, 1969). Linear-bilinear fixed effects models, such as the Sites Regression (SREG) Model (Cornelius et al., 1996; Crossa and Cornelius, 1997) and the Additive Main effect and Multiplicative Interaction (AMMI) (Gauch, 1990; Gauch and Zobel, 1996), are used for studying genotypic patterns of responses across environments. These models are parsimonious, since the interaction parameters are estimated from the singular value decomposition of the GE matrix (AMMI) or from the genotype plus genotype × environment interaction (GGE) matrix (SREG), and the patterns of response of genotypes and environments can be visualized graphically using biplots (Kempton, 1984).

For plant breeders, linear-bilinear models such as the fixed effects SREG and/or AMMI offer more opportunities for modeling GGE or GE than the simple regression of genotypes on the site mean that was previously suggested by Finlay and Wilkinson (1963) and Eberhart and Russell (1966). Several recent reviews have pointed out the merits and demerits of the fixed effect linear-bilinear models AMMI and SREG with respect to their suitability for GE analysis (Gauch, 2006; Yan and Tinker, 2006; Yan et al., 2007; Gauch et al., 2008). Yang et al. (2009) pointed out one limitation of fixed effect AMMI and SREG models, that is, these models have no inferential statistics attached to the interaction parameters used to build the biplot. Denis and Gower (1994, 1996) and Denis and Pazman (1999) proposed asymptotic confidence regions for genotypic and environmental scores that help breeders make more reliable decisions on genotype selection and recommendation; however, these confidence regions are not easily implemented

for models with more than two bilinear terms, and they require restrictive assumptions such as asymptotic normality. Yang et al. (2009) advocated the use of a non-parametric resampling technique (bootstrapping,) for constructing confidence regions for genotypic and environmental scores that can be applied to fixed effect and mixed linear-bilinear models.

From a conceptual perspective, the main distinction between fixed and random models is the non-shrunken versus the shrunken estimators of the effects where, in general, shrinkage is associated to random effects (not to fixed effects). However, Cornelius and Crossa (1999) introduced the shrinkage estimates of the linear and bilinear parameters of the fixed linear-bilinear models (e.g. AMMI, SREG, etc.) and showed that these shrinkage estimates are as good as, or better predictors, than the Best Linear Unbiased Prediction (BLUP) of the cell means of a two-way complete random model. Furthermore, the shrinkage predictors of the parameters of the fixed linear-bilinear models proposed by Cornelius and Crossa (1999) can be conceptualized as Bayesian estimates with non-informative priors.

The first proponents of the mixed linear-bilinear AMMI model for the analysis of MET were Piepho (1997, 1998), Smith et al. (2001), and Piepho and Mohring (2006). The mixed version of SREG and AMMI naturally leads to a factor analytic form for the genetic variance-covariance for environments. Since the above mentioned models are linear mixed models, they have the usual advantages in comparison with ordinary fixed effects linear-bilinear AMMI and SREG models. That is, error variance modeling can be accommodated; in particular, heterogeneity of block and error variance between environments and within-environment spatial correlation, and incomplete data are handled with ease. Furthermore, when genotypes are considered as random effects, coefficients of parentage can be incorporated into the factor analytic form for modeling GE or GGE of the mixed AMMI and SREG, respectively, hence obtaining more precise estimates of the breeding values of genotypes (Crossa et al., 2006; Oakey et al., 2006; Burgueño et al., 2007). Burgueño et al. (2008) showed how to use the factor analytic form of the mixed SREG model for clustering sites and genotypes with statistically negligible crossover interaction.

Apart from the above statistical and biological merits of the mixed linear-bilinear AMMI and SREG, it is clear that these models can deal with unbalanced GE two-way table within the analysis of MET without the need for imputing missing data as a first step and then performing

the analysis and estimating the GE parameters. However, it is unclear how asymptotic parametric confidence regions constructed under the fixed-effects model can be extended under a mixed-effects model.

Bayes' paradigm has the main feature for drawing inferences about hypotheses because it provides the flexibility of using not only observed data obtained from the current experiment but also prior information and data from previous studies; however, its use in modeling GE in the context of METs and for selecting genotypes has been very limited. A Bayesian approach to the analysis of incomplete genotype × environment × year data when using METs was presented by Theobald et al. (2002). Foucteau and Denis (2000) used a sequential Bayesian approach to make variety recommendations in the context of integration and use of information for different METs. Edwards and Jannink (2006) applied a Bayesian methodology for the analysis of GE with heterogeneous variance among environments and genotypes. Cotes et al. (2006) described a Bayesian approach for determining stable genotypes based on Shukla's (1972) stability variance.

Bayesian analysis of linear-bilinear models such as AMMI and SREG has been limited by the complex structure of the parameter space. While Bayesian computational methods such as the Markov chain Monte Carlo (MCMC) (Gilks et al., 1995; Gelman et al., 2003; Robert and Casella, 2004) led to an explosion of Bayesian methods in the 1990s, linear-bilinear models (i.e., AMMI) produce unique problems due to the orthonormal bases used in singular value decomposition. Viele and Srinivasan (2000) were the first to propose Bayesian estimation of parameters for the AMMI model using MCMC techniques through Gibbs sampling with embedded Metropolis-Hastings random walks. These authors showed that the Bayesian approach for estimating AMMI model parameters provides an easy method for dealing with unbalanced data and heterogeneity of variances, and produces, for non-informative prior distributions, shrinkage estimates of the linear (main effects) and bilinear (GE interaction) parameters similar to those proposed by Cornelius et al. (1996) and Cornelius and Crossa (1999) in the context of a two-way complete random effect model. Similar algorithms were used by Kateri et al. (2005) in the context of contingency table analysis.

Some practical and theoretical issues unresolved by Viele and Srnivasan (2000), such as whether the MCMC will always converge on the target posterior distribution, or whether AMMI bilinear terms can be estimated from the MCMC sample without violating model constraints,

were investigated in the unpublished 2001 Ph.D. thesis of G. Liu, who derived a set of conditional distributions for AMMI model parameters, such that a Gibbs sampler (without using embedded Metropolis-Hastings steps) for sampling from the joint posterior distribution could be used. This saves computing time and makes the algorithm more stable. A point discussed by simulation by Liu (unpublished data, 2001) is that while the singular values (λ) estimated from the standard least squares methods are biased upward, some shrinkage methods for estimating λ such as those described by Cornelius et al. (1996) and Cornelius and Crossa (1999) are superior to least squared estimates but tended to produce downward biased estimates so that non-zero λ's would be underestimated. Simulation results from Liu (unpublished data, 2001) showed that Bayesian-AMMI estimates of λ under certain priors of λ result in less or no upward bias; therefore the author concluded that Bayesian parameter estimation is better than the least square method and as good as or even better than shrinkage estimates of the parameters of the fixed linear-bilinear models proposed by Cornelius and Crossa (1999).

The objectives of this research were to show the use of the Bayesian paradigm to make inference on the linear and bilinear parameters of the AMMI model in terms of (i) how to develop conditional posteriors to the bilinear GE parameters of the AMMI models, and (ii) how to provide answers to questions posed to the fixed effect linear-bilinear AMMI such as how to incorporate inferential statistics into the biplot. One maize MET was employed to illustrate the use of Bayesian-AMMI consisted of 9 genotypes planted in 20 sites. Linear and bilinear parameters of the Bayesian-AMMI were estimated as usual for the Bayesian MCMC based methodology. For this data set, vague but proper prior distributions were used.

MATERIALS AND METHODS

Experimental data set

The data set used in this study (also used by Liu [unpublished data, 2001] and by Burgueño et al. [2008]) was employed to illustrate the use of the Bayesian-AMMI; it comprises 9 genotypes evaluated in 20 sites using a randomized block design with four replicates. The response variable analyzed was grain yield (in kilograms per hectare), and vague but proper prior distributions on

the fixed effects and singular values parameters were used to compute the joint posterior distribution.

Although the example used in this study is relatively small in size as compared with other METs with a larger number of genotypes or sites, it was used to illustrate the application of the Bayesian paradigm to the AMMI model. The application of the proposed Bayesian-AMMI procedure to larger MET data sets does not pose additional statistical or computing difficulties.

Statistical models

The AMMI model with Fixed Effects for Genotypes, Environments, and Genotype x Environment

In the context of a fixed effect model, the AMMI model for the average response $\bar{y}_{ij.}$ of *g* genotypes (i=1,2,…,*g*) evaluated in *e* environments (j=1,2,…,*e*) may be expressed as:

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

where μ is a general mean, α_i is the effect of the *i*th genotype, and β_j is the effect of the *j*th environment. The quantity $\lambda_k u_{ik} v_{jk}$ is the *k*th interaction component that corresponds to the *ij*th element of the singular value decomposition of the GE matrix of interactions (i.e., deviations of cell means from the additive main effects). In particular, λ_k is the singular value of the *k*th multiplicative component. These λ_k are ordered, i.e., $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_t$. The parameter u_{ik} is the *i*th element of the *k*th singular vector for rows (i.e., genotypes) and v_{jk} is the *j*th element of the *k*th singular vector for columns (i.e., environments). The u_{ik} and v_{jk} satisfy the orthonormalization constraints $\sum_i u_{ik} u_{ik'} = \sum_j v_{jk} v_{jk'} = 0$ for $k \neq k'$ and $\sum_i u_{ik}^2 = \sum_j v_{jk}^2 = 1$. The term $\bar{e}_{ij}$ is the average error associated with the response of the *i*th genotype in the *j*th environment and $t \leq \min(e-1, g-1)$.

Cornelius and Seyedsadr (1997) generalized [1], as well as other linear-bilinear models for a two-way table, into a General Linear-Bilinear Model (GLBM) of the form $\bar{y}_{ij} = \sum_{k=1}^{w} \beta_k x_{kij} + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk} + \bar{e}_{ij}$, where β_k are parameters (regression coefficients) for the linear terms, x_{kij} are known regressor variables, and λ_k, u_{ik}, v_{jk}, and $\bar{e}_{ij}$ are the same as in [1]. The number of linear (additive) terms is denoted as *w*, and the number of bilinear (multiplicative) terms is represented by *t*. In matrix notation, the previous equation can be expressed as $\mathbf{y} = \sum_{k=1}^{w} \beta_k \mathbf{X}_k + \mathbf{UDV'} + \mathbf{E}$, where $\mathbf{y} = [\bar{y}_{ij}]$, $\mathbf{X}_k = \mathbf{[x_{kij}]}$, $\mathbf{U} = [u_{ik}]$, $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_t)$, $\mathbf{V} = [v_{jk}]$, and $\mathbf{E} = [\bar{e}_{ij}]$; $\mathbf{U'U = V'V = I}_t$.

For *r* replicates, the vector **y** of *n* = *rge* phenotypic responses in the AMMI model can be expressed in matrix notation as:

$$\mathbf{y} = \mathbf{1}_n \mu + \mathbf{X}_1\boldsymbol{\alpha} + \mathbf{X}_2\boldsymbol{\beta} + \sum_{k=1}^{t} \lambda_k \text{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k + \mathbf{e} \qquad [2]$$

where $\mathbf{1}_n$ is a vector of one's of order $n \times 1$; μ is the overall mean or reference value; $\mathbf{X}_1$ is the design matrix for genotypes of order $n \times g$; $\boldsymbol{\alpha}$ is a $g \times 1$ vector of fixed genotype main effects; $\mathbf{X}_2$ and $\boldsymbol{\beta}$ are the design matrices for environments of order $n \times e$ and the $e \times 1$ vector of fixed environment main effects, respectively; λ_k is the singular value for the *k*th principal component; $\mathbf{u}_k$ and $\mathbf{v}_k$ are the *k*th singular vectors for genotypes and environments, respectively, and thus form orthonormal matrices in the singular value decomposition; and **e** is an *n*-vector of random residual effects. The vector **e** has a multivariate normal distribution with zero mean and variance-covariance matrix $\sigma_e^2\ \mathbf{I}_n$. Then, the vector of observations **y** has also a multivarite normal distribution.

The Bayesian-AMMI

Bayesian prior and posterior density of the parameters of the AMMI model

The Bayesian estimation of the parameters of the model of [2] assumes that the conditional distribution of **y**, given $\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{u}$, **v**, and σ_e^2 is a multivariate normal distribution

$$\mathbf{y} \mid \mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{u}, \mathbf{v}, \sigma_e^2 \sim N\left(\mathbf{1}_n \mu + \mathbf{X}_1\boldsymbol{\alpha} + \mathbf{X}_2\boldsymbol{\beta} + \sum_{k=1}^{t} \lambda_k \mathrm{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k, \mathbf{I}_n\sigma_e^2\right) \quad [3]$$

where $\mathbf{I}_n$ is the identity matrix of order *n*. Two criteria are used for assigning priors. First, how much information (if any) is known about the parameters, and second, that the posterior can be easily derived because it has the same distribution form as the prior (conjugacy). When nothing is known about the parameter, a non-informative or vague prior is used.

The following prior distributions, which are the same as those used by Viele and Srinivasan (2000) and Liu (unpublished data, 2001), were chosen for parameters in [3]. Subscripted symbols μ and σ^2 denote mean and variance of the prior distribution of whatever parameter is shown as subscript.

$$\mu \mid \mu_\mu, \sigma_\mu^2 \sim \mathrm{N}\left(\mu_\mu, \sigma_\mu^2\right)$$

$$\boldsymbol{\tau} \mid \boldsymbol{\mu}_\alpha, \sigma_\alpha^2 \sim N\left(\boldsymbol{\mu}_\alpha, \mathbf{I}_g\sigma_\alpha^2\right)$$

$$\boldsymbol{\delta} \mid \boldsymbol{\mu}_\beta, \sigma_\beta^2 \sim N\left(\boldsymbol{\mu}_\beta, \mathbf{I}_e\sigma_\beta^2\right)$$

$\lambda_k \mid \mu_{\lambda_k}, \sigma_{\lambda_k}^2 \sim \mathrm{N}^+\left(\mu_{\lambda_k}, \sigma_{\lambda_k}^2\right)$ with the restrictions $\lambda_k > 0$ and $\lambda_{k-1} \geq \lambda_k$

$\mathbf{u}_k \sim$ Spherical Uniform on the corrected subspace

$\mathbf{v}_k \sim$ Spherical Uniform on the correct subspace

$$\sigma_e^2 \mid \nu_e, s_e^2 \sim \mathrm{Inv-Scaled-}\chi^2\left(\nu_e, s_e^2\right)$$

where N() denotes the normal distribution, N^+ is the positive normal distribution, and $\mathrm{Inv-Scaled-}\chi^2$ is the inverse scaled chi-squared distribution. All the above priors satisfy the model constraints and are conditionally conjugate priors. The spherical uniform distribution is a special case of a von Mises-Fisher distribution (Mardia et al., 1979), which is a conjugate family. Whether the priors for the parameters are informative or non-informative depends upon their parameters; for normal priors, infinite variance makes the prior nearly non-informative.

Details of the prior spherical uniform distribution of GE parameters are in Viele and Srinivasan (2000); they refer to the constraints on the vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ that must have zero sum (as mentioned above) and unit length. The constraint that $\mathbf{u}_k$ (and $\mathbf{v}_k$) must sum to 0 is equivalent to $\mathbf{u}_k$ (and $\mathbf{v}_k$) being orthogonal to the **1** vector. Then vector $\mathbf{u}_k$ and the **1** vector form a hyperplane, and the constraint of unit length corresponds to the vectors in the unit sphere. The intersection of these subspaces is in a multi-dimensional sphere. The support of $\mathbf{u}_k$ (and $\mathbf{v}_k$) is not trivial, and Viele and Srinivasan (2000) described how to sample from the spherical uniform distributions using the correct supports.

Above, $\boldsymbol{\mu}_\alpha$ and $\mathbf{I}_g\sigma^2_\alpha$ are the prior vector of means and the prior covariance matrix of the genotypic main effects; $\boldsymbol{\mu}_\beta$ and $\mathbf{I}_e\sigma^2_\beta$ are the prior vector of means and the prior covariance matrix for the environment main effects; μ_{λ_k} and $\sigma^2_{\lambda_k}$ are the prior means and the variances for the singular value λ_k; q and σ^2_μ are the number of effects and the variance of μ, respectively; and, finally, ν_e and s^2_e are the degree of belief and the scale factor for σ^2_e. The values of μ_μ, σ^2_μ, $\boldsymbol{\mu}_\alpha$, σ^2_α, $\boldsymbol{\mu}_\beta$, σ^2_β, μ_{λ_k}, $\sigma^2_{\lambda_k}$, q, ν_e, and s^2_e are chosen with the aim of representing the prior belief about the model parameters. Furthermore, the sum to zero constraint for the elements of the $\boldsymbol{\mu}_\alpha$ and $\boldsymbol{\mu}_\beta$ vectors was considered in order to avoid confounding μ with the vectors of main effects and thereby avoid possible identifiability problems when non-informative priors are used.

Multiplying the prior distributions by the likelihood function, we obtain the following joint posterior distribution:

$$
\begin{aligned}
&p(\mu,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\lambda},\mathbf{u},\mathbf{v},\sigma_e^2 \mid \mathbf{y}) \\
&\propto \exp\left[-\frac{1}{2\sigma_\mu^2}\left(\mu_\mu-\mu\right)'\left(\mu_\mu-\mu\right)\right]\times\exp\left[-\frac{1}{2\sigma_\alpha^2}\left(\boldsymbol{\mu}_\alpha-\boldsymbol{\alpha}\right)'\left(\boldsymbol{\mu}_\alpha-\boldsymbol{\alpha}\right)\right] \\
&\times\exp\left[-\frac{1}{2\sigma_\beta^2}\left(\boldsymbol{\mu}_\beta-\boldsymbol{\beta}\right)'\left(\boldsymbol{\mu}_\beta-\boldsymbol{\beta}\right)\right]\times\prod_{k=1}^{t}\exp\left[-\frac{1}{2\sigma_\lambda^2}\left(\mu_\lambda-\lambda_k\right)'\left(\mu_\lambda-\lambda_k\right)\right] \\
&\times\exp\left[\begin{array}{l}-\frac{1}{2\sigma_e^2}\left(\mathbf{y}-\mathbf{1}_n\mu+\mathbf{X}_1\boldsymbol{\alpha}+\mathbf{X}_2\boldsymbol{\beta}+\Sigma_{k=1}^{t}\lambda_k\,\mathrm{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k\right)' \\ \left(\mathbf{y}-\mathbf{1}_n\mu+\mathbf{X}_1\boldsymbol{\alpha}+\mathbf{X}_2\boldsymbol{\beta}+\Sigma_{k=1}^{t}\lambda_k\,\mathrm{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k\right)\end{array}\right] \\
&\times\left(\left(\sigma_e^2\right)^{-\frac{n+\nu_e}{2}-1}\exp\left[-\frac{1}{2\sigma_e^2}\nu_e s_e^2\right]\right) \qquad [4]
\end{aligned}
$$

Restrictions are $\lambda_k > 0$, $\lambda_{k-1} \geq \lambda_k$, $\mathbf{u}_k'\mathbf{1} = \mathbf{v}_k'\mathbf{1} = 0$, and $\mathbf{u}_k'\mathbf{u}_{k^*} = \mathbf{v}_k'\mathbf{v}_{k^*} = 0$ (for $k \neq k^*$). In the sequel, we will refer to the methodology described here as Bayesian-AMMI.

Considerations for assigning values to the Bayesian prior distribution

When no prior information for parameters μ, α_i, β_j, and λ_k is available, in order to set priors noninformative (or nearly so), we use 0 as prior mean of all genotypic and environmental effects and large values for the prior variances. Therefore, vague but proper prior distributions on the hyperparameters were used to compute the posterior distribution. Prior hyperparameters for the (nearly) non-informative priors were chosen as $\mu_\mu = 0$, $\boldsymbol{\mu}_\alpha = \mathbf{1}_g \times 0$, $\boldsymbol{\mu}_\beta = \mathbf{1}_e \times 0$, and $\mu_{\lambda_k} = 0$ for the prior means, and equal to 1×10^8 for each of the prior variances σ_μ^2, σ_α^2, σ_β^2, and $\sigma_{\lambda_k}^2$.

For dispersion parameter σ_e^2, prior information is given by two hyperparameters: ν_e, the degree of belief (or prior degrees of freedom), and s_e^2, a scale parameter. Researchers can translate their prior knowledge or belief about the variances into numerical values of s_e^2 and ν_e

which reflect the researcher's high confidence (small variance and high values of ν_e) or lack of confidence (large variance and small values of ν_e) that parameters are equal to their respective prior mean. When the values of the hyperparameters s_e^2 and ν_e are from previous experimental data, they correspond to the estimated variance and their corresponding degrees of freedom, respectively. Both hyperparameters, ν_e and s_e^2, should be greater than zero in order to avoid obtaining an improper posterior distribution, as pointed out by Hobert and Casella (1996). The prior distribution of the error variance σ_e^2 was set at $s_e = 10^8$ with $\upsilon_e = 1$ degree of freedom.

Evaluating the posterior distribution using the Gibbs sampler algorithm

The Gibbs sampler algorithm was used to simulate from the marginal posterior distributions of the parameters of [4]. We merged μ, α_i, and β_j parameters into one vector named **θ**, and created a diagonal matrix **M** containing the prior variance of each parameter, plus the error variance (see Appendix A). After some algebraic manipulation, we found that the conditional posterior distribution for $\boldsymbol{\theta}$ has a multivariate normal distribution (see Appendix A):

$$\boldsymbol{\theta} \mid others \sim N\left(\hat{\hat{\boldsymbol{\theta}}}, \left(\mathbf{W}'\mathbf{W} + \mathbf{M}\right)^{-1} \sigma_e^2\right) \qquad [5]$$

where $\mathbf{W} = \begin{bmatrix} 1 & \mathbf{X_1} & \mathbf{X}_2 \end{bmatrix}$, and *others* indicates the observation data **y**, the other parameters in the AMMI model, except for those considered in the vector $\boldsymbol{\theta}$; $\hat{\hat{\boldsymbol{\theta}}} = \left(\mathbf{W}'\mathbf{W} + \mathbf{M}\right)^{-1}\left(\mathbf{W}'\mathbf{h}_1 + \mathbf{M}\boldsymbol{\theta}_0\right)$ and $\mathbf{h}_1 = \mathbf{y} - \sum_{k=1}^{t} \lambda_k \text{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k$.

A similar procedure was used to develop the conditional posterior of λ_k. We found that, due to restrictions in the parameter λ_k, this distribution is a degenerate multivariate normal distribution (see Appendix A):

$$\lambda_k \mid \text{others} \sim N\left(\hat{\hat{\lambda}}_k, \left(\mathbf{X}_3'\mathbf{X}_3 + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2}\right)^{-} \sigma_e^2\right) \quad \lambda_k > 0 \text{ and } \lambda_{k-1} > \lambda_k \qquad [6]$$

where

$$\hat{\hat{\lambda}}_k = \left(\mathbf{X}_3{}' \mathbf{X}_3 + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-} \left(\mathbf{X}_3{}' \mathbf{h}_2 + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \mu_{\lambda_k} \right) \text{ and } \mathbf{h}_2 = \mathbf{y} - \mathbf{W}\boldsymbol{\theta} - \sum_{k^* \neq k} \lambda_{k^*} \text{diag}(\mathbf{X}_1 \mathbf{u}_{k^*}) \mathbf{X}_2 \mathbf{v}_{k^*}$$

To improve the computational algorithm, Eq. [6] should be rewritten without a matrix notation (G. Liu, unpublished data, 2001) as follows:

$$\lambda_k \mid \text{others} \sim N\left(\frac{r\sigma_{\lambda_k}^2 \sum_{i,j} u_{ij} v_{ik} \bar{y}_{ij.} + \sigma_e^2 \mu_{\lambda_k}}{\sigma_{\lambda_k}^2 + \sigma_e^2}, \frac{\sigma_{\lambda_k}^2 \sigma_e^2}{n\sigma_{\lambda_k}^2 + \sigma_e^2} \right) \lambda_k > 0 \text{ and } \lambda_{k-1} > \lambda_k \qquad [7]$$

where *r* is the number of replicates in the experiment.

The conditional posterior distributions for $\mathbf{u}_k$ and $\mathbf{v}_k$ were given by Liu (unpublished data, 2001) as:

$$p(\mathbf{u}_k \mid others) \propto \exp\left[\frac{r\lambda_k}{\sigma_e^2} \mathbf{u}_k' \dot{\mathbf{u}}_k \right]$$

$$p(\boldsymbol{\gamma}_k \mid others) \propto \exp\left[\frac{r\lambda_k}{\sigma_e^2} \mathbf{v}_k' \dot{\mathbf{v}}_k \right]$$

where $\dot{\mathbf{u}}_k = \mathbf{Y}\mathbf{v}_k$, $\dot{\mathbf{v}}_k = \mathbf{Y}\mathbf{u}_k$, and $\mathbf{Y}$ is the $g \times e$ matrix of empirical cell means, that is, $\mathbf{Y} = [y_{ij.}]$, $i = 1, 2, \ldots, g$; and $j = 1, 2, \ldots, e$.

Liu (unpublished data, 2001) demonstrated that the restrictions $\mathbf{u}_k' \mathbf{1} = \mathbf{v}_k' \mathbf{1} = 0$, $\mathbf{u}_k' \mathbf{u}_{k^*} = \mathbf{v}_k' \mathbf{v}_{k^*} = 0 \quad k \neq k^*$ helped define, through the use of the von Mises- Fisher distribution (Mardia et al., 1979), two alternative random variables, which we denote as $\ddot{\mathbf{u}}_k$ and $\ddot{\mathbf{v}}_k$, such that:

$$p(\ddot{\mathbf{u}}_k \mid others) \propto \exp\left[\frac{rc_k\lambda_k}{\sigma_e^2} \ddot{\mathbf{u}}_k' \tilde{\mathbf{u}}_k \right] \qquad [8]$$

where $\tilde{\mathbf{u}}_k = c_k^{-1}\mathbf{H}_k'\dot{\mathbf{u}}_k$, $c_k = \sqrt{\dot{\mathbf{u}}_k'\mathbf{H}_k\mathbf{H}_k'\dot{\mathbf{u}}_k}$, and $\mathbf{H}_k'$ is a *g x (g-t)* matrix, such that its columns are orthonormal and orthogonal to vector $\mathbf{1}_g$ and the vectors $\mathbf{u}_{k^*}$ for $k^* \neq k$.

For the $\ddot{\mathbf{v}}_k$ parameter, the conditional distribution is:

$$p(\ddot{\mathbf{v}}_k \mid others) \propto \exp\left[\frac{rd_k\lambda_k}{\sigma_e^2}\ddot{\mathbf{v}}_k'\tilde{\mathbf{v}}_k\right] \qquad [9]$$

where $\tilde{\mathbf{v}}_k = d_k^{-1}\mathbf{R}_k'\dot{\mathbf{v}}_k$ and $\mathbf{R}_k'$ is an *c x (c-t)* matrix such that its columns are orthonormal and orthogonal to vector $\mathbf{1}_c$ and the other vectors $\boldsymbol{\gamma}_{k^*}$, for $k^* \neq k$, $\dot{\mathbf{v}}_k = \mathbf{Y'u}_k$, and $d_k = \sqrt{\dot{\mathbf{v}}_k'\mathbf{R}_k\mathbf{R}_k'\dot{\mathbf{v}}_k}$

.

The conditional posterior distribution for σ_e^2 is a scaled inverse χ^2 distribution

$$\sigma_e^2 \mid others \sim Inv - Scaled - \chi^2\left(\hat{\nu}_e, \hat{s}_e^2\right), \qquad [10]$$

where $\hat{\nu}_e = \nu_e + n$ and $\hat{s}_e^2 = \dfrac{\mathbf{e'e} + \nu_e s_e^2}{\nu_e + n}$.

Finally, the necessary steps to run the Gibbs sampler algorithm for the AMMI model are:

1. Set initial values for parameters σ_u^2, σ_e^2, λ_k,$\mathbf{u}_k$,$\mathbf{v}_k$;
2. Generate $\boldsymbol{\theta}$ from distribution [5] and update;
3. For each *k = 1,2,...,t*, sequentially execute the following steps a, b and c:
   a. Generate λ_k from distribution [7] and update;
   b. Generate $\mathbf{u}_k$ from distribution [8] and update;
   c. Generate $\mathbf{v}_k$ from distribution [9] and update;
4. Generate σ_e^2 from distribution [10] and update.

For the Gibbs sampler, two Markov Chains, each one of size 35000, were generated with a burn-in period of 5000 iterations. Graphical monitoring and convergence tests (Raftery and Lewis, 1995; Heidelberger and Welch, 1983; Gelman and Rubin, 1992) were used to determine when the MCMC chain reached the target distribution. After this, a thinning of size 2 for each MCMC sample was used to produce the final MCMC sample from the joint posterior distribution. Thus, a sample of size 30000 was used to estimate each marginal posterior distribution.

Bayes point estimates and Bayesian Credible Intervals

The Bayesian estimates under quadratic loss function of σ_e^2, λ_k, and $\boldsymbol{\theta}$ are the marginal posterior means, which are estimated with the MCMC sample means.

To obtain the Bayesian High Posterior Density (HPD) region at 95% probability level for σ_e^2, $\lambda_k, \mathbf{u}_k, \mathbf{v}_k$, and $\boldsymbol{\theta}$ parameters, we used the property that these intervals are the narrowest of all possible intervals at 95% (Chen and Shao, 1999). Therefore, the algorithm to obtain a 100(1-α) % HPD (0<α<1) region is as follows:

1. Sort the sample chains of the marginal posterior distribution for each parameter (σ_e^2, $\lambda_k, \mathbf{u}_k, \mathbf{v}_k$, and $\boldsymbol{\theta}$) in ascending order, i.e., from lowest to highest.

2. Calculate the length of *s* a subsample such that it contains 95% of the total samples of the marginal posterior distribution, that is, if there are 10,000 samples, then *s* = 9500.

3. Obtain the first possible interval at 95% holding the first *s*th sample values, and calculate the difference between limits.

4. Shift the reading frame on the chain in one sample and obtain the second possible interval by keeping the second and *(s + 1)* th sample values.

5. Continue this procedure until obtaining the HPD as the interval with the lowest difference between lower and upper values.

6. Repeat steps 2-5.

Note that this algorithm requires that the posterior distribution be unimodal (otherwise the HPD region could potentially be multiple disconnected intervals), which is not a problem in this context.

Convergence diagnostics for the MCMC sample

The test of Raftery and Lewis (1995) was used as a diagnostic tool for convergence where the dependence factor (I), the length of "burn in", the required sample size, and the minimum sample size based on zero autocorrelation are calculated. Estimates of the dependence factor greater than 5 (I>5) indicate strong autocorrelation, which may be due to a poor choice of starting value, high posterior correlations, or *stickiness* of the MCMC algorithm. Another diagnostic test used for convergence was that of Heidelberger and Welch (1983). The convergence diagnostic tests the null hypothesis that the sampled values come from a stationary distribution. The test is successively applied, firstly to the whole chain, then, after discarding the first 10%, 20%, etc. of the chain, until either the null hypothesis is accepted, or 50% of the chain has been discarded. The latter outcome constitutes *failure* of the stationary test and indicates that a longer MCMC run is needed. The half-width test calculates a 95% confidence interval for the mean, using the portion of the chain which passed the stationary test. Half the width of this interval is compared with the estimate of the mean. When the half-width test is not passed, the length of the sample is not long enough to estimate the mean with sufficient accuracy.

The last convergence test used was that of Gelman and Rubin (1992), which calculates the *potential scale reduction factor*, together with the upper and lower confidence limits. Approximate convergence is diagnosed when the upper limit is close to 1. For multivariate chains, a multivariate value is calculated that bounds above the potential scale reduction factor for any linear combination of the (possibly transformed) variables.

Inferential statistics of the Bayesian-AMMI biplots using the HPD regions

For each of the GE interaction factors $(\lambda_1^{1/2}u_{i1}, \lambda_2^{1/2}u_{i2})$ and $(\lambda_1^{1/2}v_{j1}, \lambda_2^{1/2}v_{j2})$ (k=1, 2), the 95% and 99% bivariate HPD regions were computed and included in the biplot graph. The HPD contours of the posterior distribution of the interaction factors delineated the regions with the highest

posterior density. Although this is not a formal statistical procedure, the HPD regions of the GE factors that include the zero on the first bilinear component axis are considered not to be statistically significantly different from zero for that component. Overlapping HPD regions from the different levels of $\mathbf{u}_k$ (or different levels of $\mathbf{v}_k$) indicate that their GE hyperparameters are not statistically significantly different between each other and therefore can be considered a part of one homogeneous group of genotypes (or environments).

Software

An R code was developed and used to perform the Bayesian-AMMI analyses. These codes can be obtained on request from any of the first three authors.

RESULTS

The three convergence diagnostics described above were run for all parameters of the Bayesian-AMMI to check that the model had good convergence properties. For all the parameters, the convergence factor was I<5, the stationary and the half-width tests were passed, and the upper confidence limit of the potential scale reduction factor was always 1 or very close to it. In addition, identifiability of all the linear parameters, the overall mean (μ), the genotypic effects ($\boldsymbol{\alpha}$), and the environmental effects ($\boldsymbol{\beta}$) for both data sets were observed throughout their correlations. All pair-wise correlations between the linear parameters of the Bayesian-AMMI model were always around 0, with some correlation values ranging from -0.2 to 0.2, indicating that complete identifiability on the estimation of the hyperparameters was achieved.

The computation of the significant tests for the bilinear terms to be retained in the traditional frequentist AMMI model, and their shrinkage estimates of the parameters of the linear and bilinear terms, can be found in Cornelius et al. (1996) and Cornelius and Crossa (1999). Liu (unpublished data, 2001) uses several Bayesian model selection criteria (Akaike´s information criterion [AIC], Bayesian information criterion [BIC], Bayes factor, and Kullback-Leibler´s information) for determining the appropriate number of bilinear terms to be retained in the final

Bayesian-AMMI model. Cornelius et al. (1996) showed that for the maize data set used in this study, two bilinear terms are significant by the F_{GH2} test. Liu (unpublished data, 2001) suggested that the number of terms to be retained in the model based on the AIC should be from 2 to 3, although AIC seemed to be liberal and choose a model with more bilinear terms. We fitted only two bilinear terms in this study for this data set.

Once a large MCMC sample is obtained from the joint posterior distribution of the parameters for the Bayesian-AMMI model with two bilinear terms, the sample mean is used to estimate the parameters. The trace and histograms of λ_1 and λ_2 showed acceptable shapes for the estimated densities (Fig. 1), with a marginal posterior distribution for λ_1 that is positive bell shaped and a marginal posterior distribution for λ_2 that moves more towards zero. For later λ's (not calculated), the distribution should be moving towards zero, as it is expected that the true values for these λ's should be closer to zero. The Bayesian-AMMI estimates of λ_1 and λ_2 were 5483 and 1627, respectively (Table 1), with a much larger SD for λ_2 (786.8) than for λ_1 (458.4); the AMMI unshrunken least squared estimates of λ_1 and λ_2 were 5923 and 3070, respectively (data not shown), and the AMMI shrunken least squared estimates of λ_1 and λ_2 obtained by the shrinkage estimate of the parameters of the fixed linear-bilinear method of Cornelius and Crossa (1999) were 5248 and 1532, respectively (data not shown).

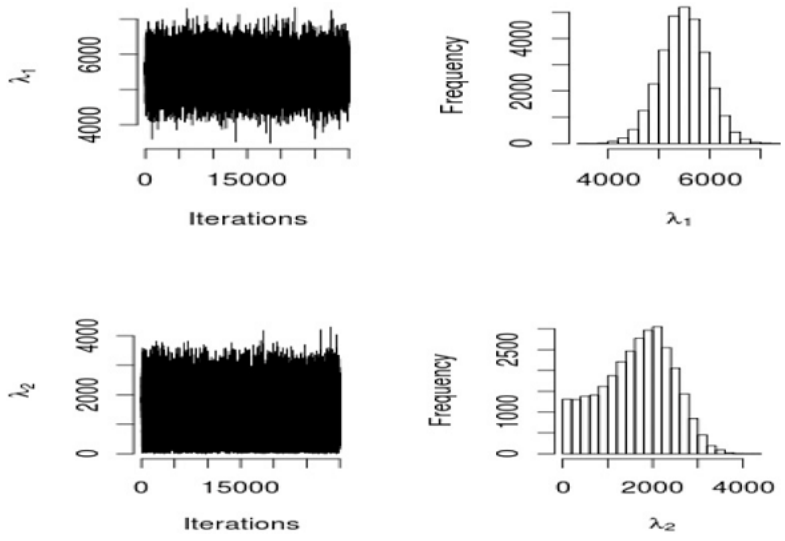


Figure 1. Traces and histograms of values of the first and second singular values (λ_1 and $\lambda_{2)}$ obtained from the Markov chain Monte Carlo (MCMC) for the maize data.

The Bayesian-AMMI estimates of the λ's are smaller than the unshrunken least squared estimates and larger than the shrinkage estimates of the λ's (though closer to the latter than the former). As expected, the Bayes estimates of the λ's rectified the upward bias of the ordinary least squares estimates and the downward bias of the shrinkage estimates.

Results of Table 1 also show the mean of the marginal posterior effects of μ, $\sigma^2_{e,}$ $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$ together with their lower and upper HPD limits. The lengths of the HPD regions of genotypes (or environments) that overlap indicate similar responses of genotypes (or environments); the posterior SD of the elements of the main effect $\boldsymbol{\alpha}$ ranged from 91.47 to 92.32, and the posterior SD o of the elements of the main effect $\boldsymbol{\beta}$ ranged from 140.8 to 143 (Table 1). The posterior means of the first and second components for the scores of sites and genotypes with their lower and upper HPD limits are in Table B1 of Appendix B, and the biplot using the mean of the scores from AMMI is depicted in Fig. 2. Some description of the response of genotypes and environments can be made in this biplot; for example, environment S8 and genotype 8 are located far from the center on the left hand side of Fig. 2, whereas genotype 4 and environment S11 are the farthest points from the center of the figure on the right hand side. However, the scores of the interaction components that are plotted in Fig. 2 are only the mode of the posterior distribution, and no statistical inference regarding similarities and dissimilarities with other genotypes and sites can be made unless HPD confidence regions around these scores using the MCMC results from the Gibbs sampler are constructed. As expected, since a vague prior was used, the Bayesian AMMI biplot is similar to that obtained from the fixed effect AMMI (data not shown). Although Burgueño et al. (2008) used the SREG, their Fig. 1 shows similar general response patterns of genotypes and sites as those depicted here in Fig. 2, especially those related to genotype 8 and site S8.

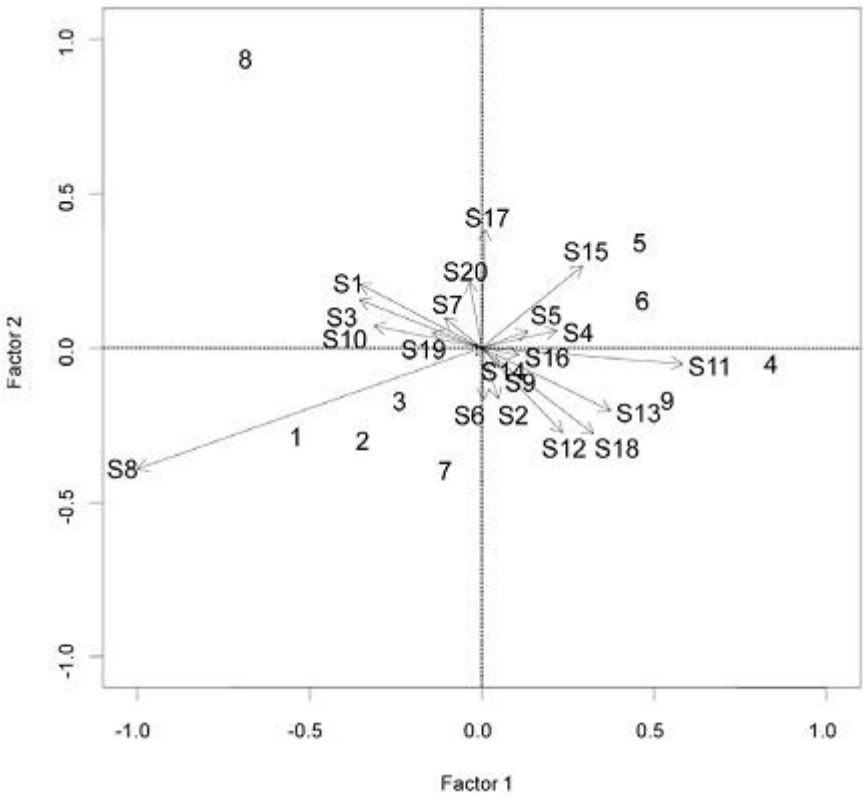


Figure 2. Biplot of Bayesian- additive main effects and multiplicative interaction (AMMI) mean posterior estimates for genotype × environment interaction components obtained from the maize multi-environment trial (MET) with nine genotypes (1-9) and 20 environments (S1-S20).

Table 1. Mean of the marginal posterior effects (mean) of the Bayesian-additive main effects and multiplicative interaction (AMMI) analysis of the maize data for grain yield (in kilograms per hectare) of nine genotypes *(i=1,2,…9)* and 20 sites *(j=1,2,…,20)* for the overall mean (μ), the dispersion parameter σ^2_e, the genotypic effects (α_i), the environmental effects (β_j), the singular values for the first and second components (λ_1 and λ_2), the corresponding standard deviation (SD), and the lower and upper highest posterior density (HPD) regions at 95%.

	Mean	SD	Lower HPD	Upper HPD
μ	4858	32.71	4792.907	4921.019
σ^2_e	763600	43580	680568.1	849750.4
α_1	-241.7	91.78	-421.8	-61.8491
α_2	-255.0	91.97	-433.048	-72.349
α_3	-34.66	91.75	-213.578	147.9183
α_4	359.2	91.68	180.6027	541.1591
α_5	389.7	92.32	206.3665	567.9183
α_6	472.5	91.47	296.9934	655.4933
α_7	-186.0	92.19	-363.985	-3.5244
α_8	-574.9	91.69	-758.735	-401.037
α_9	70.95	91.90	-107.789	247.4022
β_1	-1245	141.8	-1518.78	-963.865
β_2	-647.3	141.5	-925.309	-373.425
β_3	246.6	141.6	-28.9429	525.8115
β_4	371.6	141.6	93.7568	647.7566
β_5	96.01	141.6	-186.891	365.2221
β_6	1448	141.6	1170.011	1722.119
β_7	-1629	142.4	-1906.09	-1351.23
β_8	-830.6	142.5	-1111.01	-552.443
β_9	111.7	141.8	-164.931	391.3228
β_{10}	-1921	141.8	-2191.5	-1636.95
β_{11}	449.2	143.0	176.6458	738.1901
β_{12}	2659	141.9	2387.132	2942.396
β_{13}	1474	142.6	1188.188	1748.386
β_{14}	1194	141.8	914.8787	1470.252
β_{15}	189.3	141.6	-95.914	457.7014
β_{16}	549.3	142.4	265.906	821.7052
β_{17}	17.62	141.7	-265.469	290.3307
β_{18}	-305.5	142.0	-588.161	-34.3383
β_{19}	-2207	141.4	-2481.32	-1928.51
β_{20}	-19.72	140.8	-289.285	261.169
λ_1	5483	458.4	4573.825	6366.038
λ_2	1627	786.8	2.0887	2822.725

Uncertainty and confidence regions of the first two bilinear terms of the Bayesian-AMMI

In general it can be seen from Table B1 (Appendix B) that values of u_{i1} and v_{j1} are larger than their corresponding values of u_{i2} and v_{j2}, and that the SDs were much smaller and the length of the HPD regions were much narrower for the first bilinear components of genotypes and sites (u_{i1} and v_{j1}) than for their second bilinear components, u_{i2} and v_{j2}. The majority of u_{i2} and v_{j2} were two to four times smaller than u_{i1} and v_{j1}, and all the HPD regions covered the zero value. The Bayesian-AMMI second bilinear components for genotypes and sites were estimated with more uncertainty than the first bilinear components. These results indicated that the first component discriminated more the genotypes and environments than the second component and this is done with more precision in the first component than in the second component. This is also reflected in the HPD confidence regions around these scores as shown in Figs. 3 and 4. In general, this is in agreement with the theory of principal component analysis where the first component discriminates more the observations than the second and other components; the first component (that accounts for the maximum variability) supposes to have less uncertainty than the second and further components.

Figs. 3 and 4 depicted the HPD confidence regions at the 95% (internal gray contour) and 99% (internal gray contour) probability levels for some genotypes and sites, respectively. The only genotypes and sites included in these figures are those which have a contour at the 95% probability level that did not include zero for the first component, that is, those genotypes and sites with $u_{i1} = v_{j1} \neq 0$.

Concerning genotypes, it is clear that genotypes 8 and 1 on the left hand side of Fig. 3 formed a genotype group with negative values for the first bilinear term ($u_{1,1} = -0.3724$ and $u_{8,1} = -0.4661$, Table B1) that were significantly different from zero. On the right hand side of Fig. 3, genotypes 4, 5, 6, and 9 formed another distinct group of genotypes with positive values of $u_{i,1}$ ($u_{4,1} = 0.5016$, $u_{5,1} = 0.2756$, $u_{6,1} = 0.2649$, $u_{9,1} = 0.3135$, Table B1) that were significantly different from zero and also significantly different from those values of genotypes 1

and 8 on the first bilinear term ($u_{1,1}$ and $u_{8,1}$). The remaining genotypes formed another group not different from zero for the first bilinear component. The large uncertainty (high SD and wide HPD; Table B1) of the elements of $u_{i,2}$ is reflected in the long shadow areas along the axis of the second bilinear term.

Similar results for the elements of $v_{j,1}$ and $v_{j,2}$ were found in Fig. 4 for sites. Site S8 on the left hand side of Fig. 4 formed one clear group ($v_{8,1}$ = -0.6353, Table B1), and sites S11, S13, and S18 formed another set ($v_{11,1}$ = 0.3673, $v_{13,1}$ = 0.2358, $v_{18,1}$ = 0.2042, Table B1), both groups of sites had long ellipses along the second bilinear term, reflecting their large SD and wide length of the HPD as they reflect the uncertainty of these estimates. The rest of the sites showed that their first bilinear component did not differ from zero at the 99% probability level.

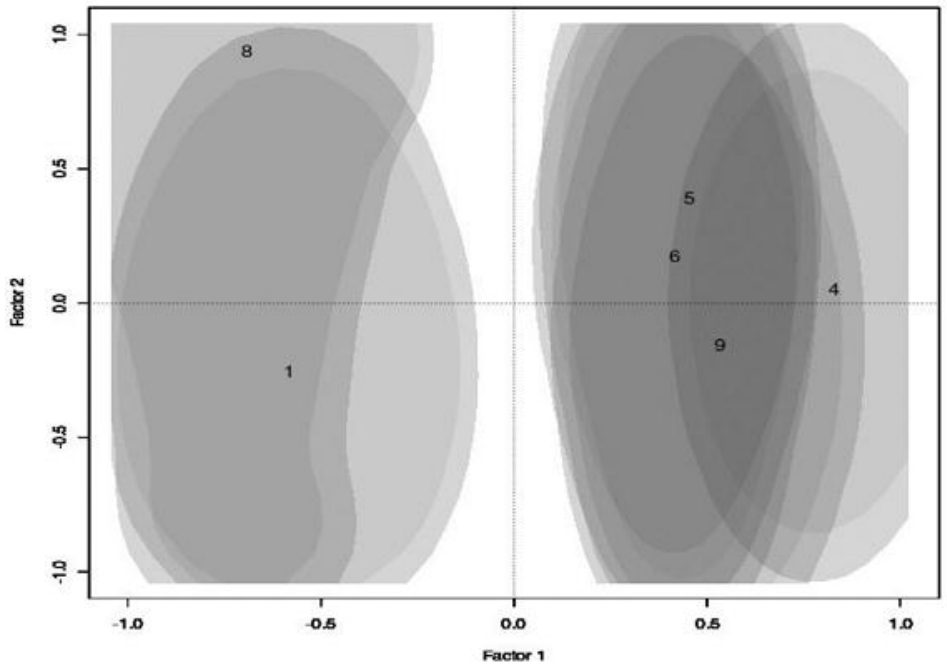


Figure 3. Biplot of Bayesian-additive main effects and multiplicative interaction (AMMI) of the mean posterior estimates for genotype × environment interaction components for genotypes 1, 4, 5, 6, 8, and 9 at the 95% (gray internal contour) and 99% (gray external contour) highest posterior density (HPD) regions obtained from the maize multi-environment (MET). Only those genotypes (1, 4, 5, 6, 8, and 9) that do not include the value of zero for the first bilinear component at the 95% probability level are depicted.

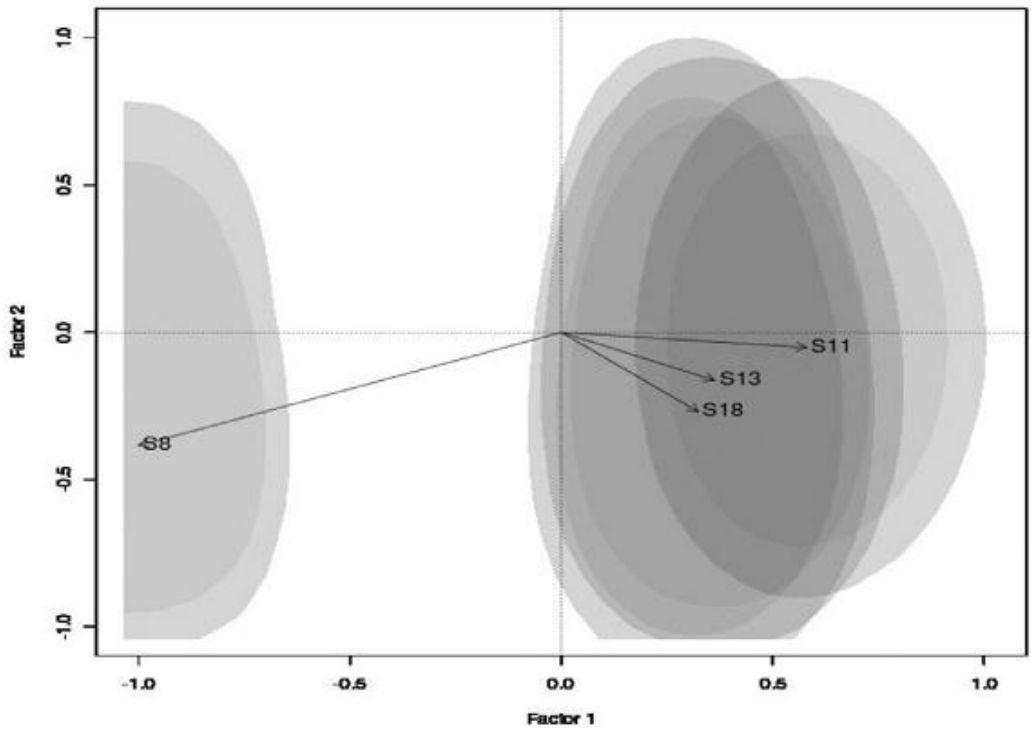


Figure 4. Biplot of Bayesian-additive main effects and multiplicative interaction (AMMI) of the mean posterior estimates for genotype × environment interaction components for environments S8, S11, S13, and S18 at the 95% (gray internal contour) and 99% (gray external contour) highest posterior density (HPD) regions obtained from the maize multi-environment (MET). Only those environments (S8, S11, S13, and S18) that do not include the value of zero for the first bilinear component at the 95% probability level are depicted.

DISCUSSION

There are examples in the literature showing that one could think of the linear-bilinear mixed AMMI model as being Bayesian in the sense that certain sets of effects have priors, and the choices for these are based on quantitative genetics, biology, and experience. For example, in the Smith et al. (2001) approach, the prior density for the genotypic effects in different trials has a factor analytic variance structure, and the prior density for the error effects has a separate spatial covariance structure within each environment. This approach has become widely accepted and used (Crossa et al., 2006; Burgueño et al., 2007, 2008; Matthews et al., 2007). In terms of within-trial prediction, Cornelius and Crossa (1999) showed that, as predictors, estimates of the parameter of the fixed linear-bilinear models are as good as, or better, than the BLUP. The shrinkage predictors of the parameters of the fixed linear-bilinear models proposed by Cornelius and Crossa (1999) can be conceptualized as Bayesian estimates with non-informative priors. However, the prediction assessment scheme of Cornelius and Crossa (1999) is based solely on predicting the performance of an individual genotype replicated within environments, and not on predicting the performance of a genotype that is completely missing in one environment.

Although the objective of this study was not to compare the predictive ability of Bayesian-AMMI versus the mixed AMMI, some preliminary results using different types of mixed AMMI (data not shown) showed that one single model is not the best across all data sets. The difficulty when attempting to compare the prediction assessment of Bayesian-AMMI with prior information versus mixed AMMI is the appropriate mixed AMMI to be used that will resemble, as much as possible, the Bayesian-AMMI.

The objective of this research was to propose a Bayesian approach to linear-bilinear AMMI that can be useful for researchers in general. Some advantages of the conditional posterior estimates of the Bayesian-AMMI model over the fixed effect linear-bilinear AMMI can be mentioned: (i) since all effects in Bayesian models are random, the Bayesian-AMMI does avoid the consistent problem of determining which effects should be fixed and which random; however, the choice of a completely noninformative prior is somewhat analogous to fitting an effect as fixed effect in the traditional frequentist paradigm, whereas an informative prior

produces information borrowing and it can be considered analogous to the case of random effect in mixed models in which there is shrinkage estimation of the parameters of the model; (ii) it provides a natural method for deriving confidence regions around the genotypic and environmental GE parameters given by their scores that allow identifying groups of similar genotypes and environments; as pointed out by Yang et al. (2009); this is an important drawback of the fixed and mixed AMMI that offers no statistical inference for identifying separable groups; (iii) it deals with the unbalanced data that are always present in MET in a natural manner; this is not a problem for the mixed AMMI but a difficult task for the fixed AMMI, unless some data imputation technique is used; although the data sets used in this study did not have missing data, the Bayesian imputation has the advantage of being part of the entire analysis, whereas the earlier approaches need to impute the missing values first and then perform the AMMI analysis, and (iv) the Bayesian-AMMI can be efficiently used to incorporate information from existing historical MET (prior) on environmental and genotypic means and dispersion parameters such as environmental, genotypic, or error variances; this information cannot be efficiently incorporated into the fixed or mixed AMMI.

Linear-bilinear models such as AMMI offer a family of models, rather than a single model, because a researcher may use AMMI with several components, up to the full model. This manuscript focuses on the Bayesian-AMMI model with two components (however, Bayesian-AMMI models with three or more components can be computed for any data set). The comparison between biplots from Bayesian-AMMI based on two yr data versus biplots obtained from fixed effect AMMI based on the second year data will depend on the correlation between years.

We recognize that the example used in this paper is quite small. Plant breeding METs are usually larger than the example used here. For example, Matthews et al. (2007) fit the random effects AMMI model to MET data with 106 environments and 41 varieties, and Thompson et al. (2003) fit the model for 62 environments and 216 varieties. Nevertheless, the example used here illustrates the use of the Bayesian-AMMI for studying the response patterns of genotypes using proper but diffuse priors. This new approach offers new opportunities for efficiently incorporating historical data on environments and genotypes that should be useful for achieving breeders' objectives as well as offering density regions around the estimated GE parameters.

Although the computer time for processing large METs can be substantially greater than that needed to fit linear-bilinear mixed AMMI, the continuous increase in computer power will minimize this disadvantage of the Bayesian-AMMI.

Although this study is a step in the right direction for applying Bayesian inference to linear-bilinear models, several topics remain to be studied and clarified. One issue is employing a formal Bayesian inference using an informative or partially informative prior distribution for all parameters of the model. Another issue that needs to be examined is the possibility of using a multivariate matrix approach to the singular value decomposition of the GE interaction parameters, rather than a vector approach as the one used here, while conserving the necessary orthonormality constraints on the bilinear hyperparameters; this would simplify the computation of the Gibbs sampler.

The present Bayes inference methodology can be extended to other linear-bilinear models such as the Sites Regression (SREG), the Genotype Regression (GREG), and the Complete Multiplicative Model (COMM); however, further research is required on this topic. Incorporation of the coefficient of parentage into the prior density for the genotypic effects in the Bayesian-AMMI should not pose any difficulty and should be a subject of further research.

REFERENCES

Burgueño, J., J. Crossa, P.L. Cornelius, R. Trethowan, G. McLaren, and A. Krishnamachari. 2007. Modeling Additive × Environment and Additive × Additive × Environment Using Genetic Covariances of Relatives of Wheat Genotypes. Crop Sci. 47:311–320.

Burgueño, J., J. Crossa, P.L. Cornelius, and R.-C. Yang. 2008. Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. Crop Sci. 48:1291-1305.

Chen, M.H., and Q.M. Shao. 1999. Monte Carlo Estimation of Bayes Credible and HPD Interval. Journal of Computational and Graphical Statistics 8:69-92.

Cornelius, P.L., J. Crossa, and M.S. Seyedsadr. 1996. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. Pp. 199-234 in Genotype-by-Environment Interaction, edited by M.S. Kang and H.G. Gauch. Boca Raton, FL: CRC press.

Cornelius, P.L., and M.S. Seyedsadr. 1997. Estimation of general linear-bilinear models for two-way tables. Journal of Statistical Computation and Simulation 58:287-322.

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.

Cotes, J.M., Crossa, J., Sanches, A., and Cornelius, P.L. 2006. A Bayesian approach for assessing the stability of genotype. Crop Sci. 46:2654-2665.

Crossa, J., and P.L. Cornelius. 1997. Sites Regression and Shifted Multiplicative models clustering of cultivar trials sites under heterogeneity of error variance. Crop Sci.. 37: 406-415.

Crossa, J., Yang, R.-C., and Cornelius, P.L. 2004. Studying cross-over genotype x environment interaction using linear-bilinear models and mixed models. Journal of Agriculture, Biology and Environmental Statistics 9:362-380.

Crossa, J., J. Burgueno, P.L. Cornelius, G. McLaren, R. Trethowan, and A. Krishnamachari. 2006. Modeling genotype × environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. Crop Sci. 46:1722-1733.

Denis, J.-B., and J.C. Gower. 1994. Asymptotic covariances for parameters of biadditve models. Util. Math. 46: 193-205.

Denis, J.-B., and J.C. Gower. 1996. Asymptotic confidence regions for biadditive models: Interpreting genotype-environment interactions. Appl. Statist. 45: 479-493.

Denis, J.-B., and A. Pazman. 1999. Bias of LS estimators in nonlinear regression models with constraints. Part II: Biadditive models. Appl. Math. 44: 375-403.

Eberhart, S.A., and Russell, W.A. 1966. Stability parameters for comparing varieties. Crop Sci. 6:36-40.

Edwards, J.W., and Jannink, J.-L. 2006. Bayesian modeling of heterogeneous error and genotype × environment interaction variances. Crop Sci. 46:820-833.

Finlay, K.W., and Wilkinson, G.N. 1963. The analysis of adaptation in a plant breeding programme. Australian Journal of Agricultural Research 14:742-754.

Foucteau, V., and Denis, J.-B. 2000. Statistical analysis of successive series of experiments in plant breeding: a Bayesian approach. Pp. 49-56 in Eucarpia, Quantitative Genetics and Breeding Methods: The Way Ahead, edited by A. Gallais, C. Dillmann, and I. Goldringer. Versailles, France: INRA Editions.

Gauch, H.G. 1990. Using interaction to improve yield estimates. Pp. 141-150 in Genotype-by-Environment Interaction and Plant Breeding, edited by M.S. Kang. Baton Rouge, LA: Department of Agronomy, Louisiana State University.

Gauch, H.G., and Zobel, R.W. 1996. AMMI analysis of yield trials. Pp. 85-122 in Genotype-by-Environment Interaction, edited by M.S. Kang and H.G. Gauch. Boca Raton, FL: CRC press.

Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 46:1488-1500.

Gauch, H.G., H.-P. Piepho, and P. Annicchiarico. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. Crop Sci. 48:866-889.

Gelman, A., and D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457-511.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 2003. Bayesian Data Analysis. Boca Raton, FL: CRC Press.

Gilks, W., Richardson, S., and Speigelhalter, D.J. 1995. Markov Chain Monte Carlo in Practice. Great Britain: Chapman and Hall.

Heidelberger, P., and P.D. Welch. 1983. Simulation run length control in the presence of an initial transient. Opns Res. 31:1109-1144.

Hobert, J.P., and Casella, G. 1996. The effect of improper priors on Gibbs Sampling in hierarchical linear mixed models. Journal of the American Statistical Association 91:1461-1473.

Kateri, M., Nicolaou, A., and Ntzoufras, I. 2005. Bayesian Inference for the RC(m) Association Model. Journal of Computational and Graphical Statistics 14: 116-138.

Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. J. Agric. Sci. 103: 123-135.

Mandel, J. 1961. Non-additivity in two-way analysis of variance. Journal of the American Statistical Association 56:878-888.

Mandel, J. 1969. The partitioning of interactions in analysis of variance. Journal of Research of the National Bureau of Standards – B. 73:309-328.

Mardia, K.V., Kent, J.T., and Bibbi, J.B. 1979. Multivariate analysis. Academic Press: London, UK.

Matthews, K.L., Chapman, S.C., Trethowan, R., Pfeiffer, W., van Ginkel, M., Crossa, J., Payne, T., DeLacy, I., Fox, P.N. and Cooper, M. 2007. Global adaptation patterns of Australian and CIMMYT spring bread wheat. Theoretical and Applied Genetics 115: 819-835.

Oakey, H., A. Verbyla, W. Pitchford, B. Cullis, and H. Kuchel. 2006. Joint modeling of additive and non-additive genetic line effects in single field trials. Theoretical and Applied Genetics 113:809-819.

Piepho, H.-P. 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. Biometrics 53: 761-767.

Piepho, H.-P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor analytic variance-covariance structures. Theoretical and Applied Genetics 97: 195-201.

Piepho, H.-P. and Mohring, J. 2006. Selection in cultivar trials - is it ignorable? Crop Sci. 46: 192-201.

Liu, G. 2001. Bayesian computation for general linear-bilinear models. Ph D. dissertation. Department of Statistics. University of Kentucky. Lexington, KY.

Raftery, A.E. and Lewis, S.M. 1995. The number of iterations, convergence diagnostics and generic Metropolis algorithms. *In* Practical Markov Chain Monte Carlo, W.R. Gilks, D.J. Spiegelhalter and S. Richardson, eds. London, U.K.: Chapman and Hall.

Robert, C., and Casella, G. 2004. Monte Carlo Statistical Methods. New York, NY: Springer-Verlag.

Shukla, G.K. 1972. Some statistical aspects of partitioning genotype-environment components of variability. Heredity 29:237-245.

Smith, A.B., Cullis, B.R., and Thompson, R. 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57: 1138-1147.

Theobald, C., Talbot, M., and Nabugoomu, F. 2002. A Bayesian approach to regional and local-area prediction from crop variety trials. Journal of Agricultural, Biological and Environmental Statistics 7:403-419.

Thompson, R., Cullis, B.R., Smith, A.B., and Gilmour, A.R. 2003. A sparse implementation of the Average Information algorithm for factor analytic and reduced rank variance models. Australian and New Zealand Journal of Statistics 45: 445-460.

Viele, K., and Srinivasan, C. 2000. Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback-Leibler information. Journal of Statistical Planning and Inference 84:2-1-219.

Yan, W., and N.A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Can. J. Plant Sci. 86:623-645.

Yan, W., M.S. Kang, B. Ma, S. Woods, and P.L. Cornelius. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Sci. 47:643-655.

Yang, R.-C., J. Crossa, P.L. Cornelius, and J. Burgueño. 2009. Biplot analysis of genotype × environment interaction: Proceed with caution. Crop Sci. 49:1564–1576.

APPENDIX A

Reparameterization of the joint posterior distribution

Vectors of the joint posterior distribution [4] are simulated and sampled using the Gibbs sampler algorithm. We consider the following reparameterization:

$$\boldsymbol{\theta}=\begin{bmatrix}\mu\\ \boldsymbol{\alpha}\\ \boldsymbol{\beta}\end{bmatrix},\ \boldsymbol{\theta}_0=\begin{bmatrix}\mu_\mu\\ \mu_{\boldsymbol{\alpha}}\\ \mu_{\boldsymbol{\beta}}\end{bmatrix},\ \mathbf{M}=\begin{bmatrix}\frac{\sigma_e^2}{\sigma_\mu^2} & & \\ & \mathbf{I}_g\frac{\sigma_e^2}{\sigma_{\boldsymbol{\alpha}}^2} & \\ & & \mathbf{I}_e\frac{\sigma_e^2}{\sigma_{\boldsymbol{\beta}}^2}\end{bmatrix}$$

$\mathbf{W}=\begin{bmatrix}1 & \mathbf{X_1} & \mathbf{X}_2\end{bmatrix}$ and $\mathbf{h}_1=\mathbf{y}-\sum_{k=1}^{t}\lambda_k\text{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k$.

Therefore, the joint posterior distribution [4] can be written as:

$$p(\mu,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\lambda},\mathbf{u},\mathbf{v},\sigma_u^2,\sigma_e^2\,|\,\mathbf{h}_1)$$
$$\propto\times\exp\left[-\frac{1}{2\sigma_e^2}\left(\mathbf{h}_1-\mathbf{W}\boldsymbol{\theta}\right)'\left(\mathbf{h}_1-\mathbf{W}\boldsymbol{\theta}\right)\right]\times\exp\left[-\frac{1}{2\sigma_e^2}\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)'\mathbf{M}\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)\right] \qquad [4a]$$
$$\times\left(\left(\sigma_e^2\right)^{-\frac{n+v_e}{2}-1}\exp\left[-\frac{1}{2\sigma_e^2}v_e s_e^2\right]\right)\prod_{k=1}^{t}\exp\left[\frac{1}{2\sigma_{\lambda_k}^2}\left(\mu_{\lambda_k}-\lambda_k\right)'\left(\mu_{\lambda_k}-\lambda_k\right)\right]$$

Conditional distribution for μ, α_i, and β_j parameters

The quadratic form $\left(\mathbf{h}_1-\mathbf{W}\boldsymbol{\theta}\right)'\left(\mathbf{h}_1-\mathbf{W}\boldsymbol{\theta}\right)+\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)'\mathbf{M}\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)=\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)'\mathbf{W}'\mathbf{W}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)+\left(\mathbf{h}_1-\mathbf{W}\hat{\boldsymbol{\theta}}\right)'\left(\mathbf{h}_1-\mathbf{W}\hat{\boldsymbol{\theta}}\right)+\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)'\mathbf{M}\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)$, where $\hat{\boldsymbol{\theta}}=\left(\mathbf{W}'\mathbf{W}\right)^{-}\mathbf{W}'\mathbf{h}_1$ and $\left(\mathbf{W}'\mathbf{W}\right)^{-}$ is the generalized inverse of $\mathbf{W}'\mathbf{W}$. The term $\left(\mathbf{h}_1-\mathbf{W}\hat{\boldsymbol{\theta}}\right)'\left(\mathbf{h}_1-\mathbf{W}\hat{\boldsymbol{\theta}}\right)$ is not dependent on $\boldsymbol{\theta}$, which can be absorbed by the proportional constant. Then, keeping the other terms constant,

$\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)'\mathbf{W}'\mathbf{W}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)$ + $\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)'\mathbf{M}\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}\right)$ = $\left(\boldsymbol{\theta}-\hat{\hat{\boldsymbol{\theta}}}\right)'\left(\mathbf{W}'\mathbf{W}+\mathbf{M}\right)\left(\boldsymbol{\theta}-\hat{\hat{\boldsymbol{\theta}}}\right)$ + $\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)'\left(\mathbf{W}'\mathbf{W}\right)\left(\mathbf{W}'\mathbf{W}+\mathbf{M}\right)^{-}\mathbf{M}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)$, where $\hat{\hat{\boldsymbol{\theta}}}=\left(\mathbf{W}'\mathbf{W}+\mathbf{M}\right)^{-}\left(\mathbf{W}'\mathbf{h}_1+\mathbf{M}\boldsymbol{\theta}_0\right)$. Because $\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)'\left(\mathbf{W}'\mathbf{W}\right)\left(\mathbf{W}'\mathbf{W}+\mathbf{M}\right)^{-}\mathbf{M}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right)$ is not dependent on **θ**, the conditional distribution of **θ** given σ_e^2 should be expressed as:

$$p(\boldsymbol{\theta}\,|\,\sigma_e^2,\mathbf{h}_1)\propto\exp\left[-\frac{1}{2\sigma_e^2}\left(\boldsymbol{\theta}-\hat{\hat{\boldsymbol{\theta}}}\right)'\left(\mathbf{W}'\mathbf{W}+\mathbf{M}\right)\left(\boldsymbol{\theta}-\hat{\hat{\boldsymbol{\theta}}}\right)\right].$$

Conditional posterior distribution for λ_k

For the conditional posterior distribution of λ_k , we considered the vector $\mathbf{h}_2=\mathbf{y}-\mathbf{W}\boldsymbol{\theta}-\sum_{k^*\neq k}\lambda_{k^*}diag(\mathbf{X}_1\mathbf{u}_{k^*})\mathbf{X}_2\mathbf{v}_{k^*}$. Then the posterior can be expressed as:

$$p(\lambda_k\,|\,\sigma_u^2,\sigma_e^2,\mathbf{h}_2)\propto\exp\left[\frac{1}{2\sigma_e^2}\left(\mathbf{h}_2-\lambda_k\mathrm{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k\right)'\left(\mathbf{h}_2-\lambda_k\mathrm{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k\right)\right]$$
$$\times\exp\left[\frac{1}{2\sigma_e^2}\left(\mu_{\lambda_k}-\lambda_k\right)'\frac{\sigma_e^2}{\sigma_{\lambda_k}^2}\left(\mu_{\lambda_k}-\lambda_k\right)\right],$$

with the restrictions $\lambda_k>0$ and $\lambda_{k-1}>\lambda_k$. Now, defining $\mathbf{X}_3=\mathrm{diag}(\mathbf{X}_1\mathbf{u}_k)\mathbf{X}_2\mathbf{v}_k$, the posterior becomes

$$p(\lambda_k\,|\,\sigma_u^2,\sigma_e^2,\mathbf{h}_2)\propto\exp\left[\frac{1}{2\sigma_e^2}\left(\mathbf{h}_2-\mathbf{X}_3\lambda_k\right)'\left(\mathbf{h}_2-\mathbf{X}_3\lambda_k\right)\right]\times\exp\left[\frac{1}{2\sigma_e^2}\left(\mu_{\lambda_k}-\lambda_k\right)'\frac{\sigma_e^2}{\sigma_{\lambda_k}^2}\left(\mu_{\lambda_k}-\lambda_k\right)\right].$$

Holding the quadratic form and using a similar algebraic procedure as used for the **θ** parameter, we can express the posterior conditional distribution for λ_k as:

$$p(\lambda_k \mid \sigma_e^2, \mathbf{h}_2) \propto \exp\left[\frac{1}{2\sigma_e^2} \left(\lambda_k - \hat{\hat{\lambda}}_k \right)' \left(\mathbf{X}_3'\mathbf{X}_3 + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right) \left(\lambda_k - \hat{\hat{\lambda}}_k \right) \right]$$

where $\hat{\hat{\lambda}}_k = \left(\mathbf{X}_3'\mathbf{X}_3 + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-} \left(\mathbf{X}_3'\mathbf{h}_2 + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \mu_{\lambda_k} \right)$.

APPENDIX B

Table B1. Values of the mean of the marginal posterior effects for the genotypic scores of the first and second components, $u_{i,1,}$ and $u_{i,2}$, respectively, and of the environmental scores of the first and second components $v_{i,1}$, and $v_{i,1}$, respectively, and the corresponding SD, and lower and upper highest posterior density (HPD) regions at 95% of the Bayesian-additive main effects and multiplicative interaction (AMMI) analysis of the maize trial of grain yield (in kilograms per hectare) with nine genotypes (i = *1,2,…9*) and 20 sites (j = *1,2,…,20*).

γ_{j1} and α_{i1}	Mean	SD	HPD lower	HPD upper	γ_{j2} and α_{i2}	Mean	SD	HPD lower	HPD upper
$v_{1,1}$	-0.2110	0.0751	-0.359	-0.0649	$v_{1,2}$	0.0840	0.2205	-0.3702	0.4797
$v_{2,1}$	0.0169	0.0757	-0.129	0.1685	$v_{2,2}$	-0.0562	0.2060	-0.4417	0.3544
$v_{3,1}$	-0.2427	0.0752	-0.3911	-0.0955	$v_{3,2}$	0.0539	0.2187	-0.3744	0.4739
$v_{4,1}$	0.1389	0.0734	-0.0066	0.2807	$v_{4,2}$	0.0271	0.1869	-0.3382	0.3944
$v_{5,1}$	0.0811	0.0754	-0.0605	0.2365	$v_{5,2}$	0.0300	0.1952	-0.3524	0.4063
$v_{6,1}$	-0.0068	0.0766	-0.1572	0.1416	$v_{6,2}$	-0.0621	0.2177	-0.4581	0.3876
$v_{7,1}$	-0.0763	0.0753	-0.2203	0.0718	$v_{7,2}$	0.0331	0.1974	-0.3635	0.4119
$v_{8,1}$	-0.6353	0.0613	-0.7520	-0.5125	$v_{8,2}$	-0.1302	0.2395	-0.5169	0.3582
$v_{9,1}$	0.0359	0.0752	-0.1071	0.1869	$v_{9,2}$	-0.0173	0.1919	-0.3932	0.3645
$v_{10,1}$	-0.2052	0.0729	-0.3499	-0.0648	$v_{10,2}$	0.0184	0.1827	-0.3433	0.3768
$v_{11,1}$	0.3673	0.0690	0.2342	0.5038	$v_{11,2}$	-0.0189	0.1773	-0.3680	0.3237
$v_{12,1}$	0.1510	0.0754	0.0061	0.3009	$v_{12,2}$	-0.0967	0.2279	-0.4928	0.3717
$v_{13,1}$	0.2358	0.0761	0.0860	0.3849	$v_{13,2}$	-0.0762	0.2350	-0.4971	0.3989
$v_{14,1}$	-0.0058	0.0743	-0.1514	0.1406	$v_{14,2}$	-0.0074	0.1795	-0.3583	0.3467
$v_{15,1}$	0.1950	0.0755	0.0480	0.3428	$v_{15,2}$	0.0884	0.2229	-0.3642	0.4810
$v_{16,1}$	0.0746	0.0739	-0.0681	0.2205	$v_{16,2}$	-0.0014	0.1822	-0.3671	0.3533
$v_{17,1}$	-0.0008	0.0783	-0.1526	0.1525	$v_{17,2}$	0.1313	0.2619	-0.3998	0.5649
$v_{18,1}$	0.2042	0.0769	0.0552	0.3562	$v_{18,2}$	-0.1026	0.2524	-0.5340	0.4198
$v_{19,1}$	-0.0826	0.0761	-0.2261	0.0709	$v_{19,2}$	0.0246	0.2145	-0.3929	0.4444
$v_{20,1}$	-0.0343	0.0760	-0.1831	0.1130	$v_{20,2}$	0.0784	0.2167	-0.3544	0.4799
$u_{1,1}$	-0.3724	0.0742	-0.5197	-0.2286	$u_{1,2}$	-0.0841	0.3087	-0.6265	0.5484
$u_{2,1}$	-0.2416	0.0756	-0.3910	-0.0972	$u_{2,2}$	-0.1145	0.2964	-0.6216	0.4926
$u_{3,1}$	-0.1603	0.0753	-0.3089	-0.0153	$u_{3,2}$	-0.0568	0.262	-0.5666	0.4727
$u_{4,1}$	0.5016	0.0647	0.3761	0.6277	$u_{4,2}$	0.0059	0.2222	-0.4318	0.4545
$u_{5,1}$	0.2756	0.0775	0.1229	0.4253	$u_{5,2}$	0.1241	0.3260	-0.5488	0.6650
$u_{6,1}$	0.2649	0.0727	0.1194	0.4043	$u_{6,2}$	0.0820	0.2619	-0.4638	0.5546
$u_{7,1}$	-0.1152	0.0827	-0.2744	0.0503	$u_{7,2}$	-0.1333	0.3656	-0.7430	0.6020
$u_{8,1}$	-0.4661	0.0776	-0.6135	-0.3118	$u_{8,2}$	0.2483	0.4262	-0.6113	0.7947
$u_{9,1}$	0.3135	0.0741	0.1712	0.4608	$u_{9,2}$	-0.0716	0.2805	-0.5972	0.4905

CAPÍTULO 2: A GENERAL BAYESIAN ESTIMATION METHOD OF LINEAR-BILINEAR MODELS APPLIED TO PLANT BREEDING TRIALS WITH GENOTYPE × ENVIRONMENT INTERACTION

Diego JARQUIN, Sergio PEREZ-ELIZALDE, and Jose CROSSA

RESUMEN

El análisis estadístico de tablas de doble entrada con interacción surge en diferentes campos de investigación. Este estudio propone la distribución von Mises-Fisher como a priori en el conjunto de matrices ortogonales en un modelo lineal-bilineal para el estudio e interpretación de la interacción de una tabla de doble entrada. Datos simulados y datos empíricos de un ensayo de fitomejoramiento fueron usados para ilustración; los datos empíricos consisten de un experimento de múltiples ambientes establecido en dos años consecutivos. Para los datos simulados, se utilizaron distribuciones a priori vagas pero propias, y para los datos reales de fitomejoramiento, las observaciones del primer año fueron usadas como información a priori para los parámetros del modelo en el análisis del segundo año. Regiones bivariadas de alta densidad a posterior (HPD) para los scores a posteriori son mostrados en los biplots y la significancia de los términos bilineales se determino mediante el factor de Bayes. Los resultados del experimento de fitomejoramiento muestran la utilidad de esta aproximación general bayesiana en ensayos de mejoramiento para la detección de grupos de genotipos y de ambientes que causan una interacción (genotipo x ambiente) significativa. La metodología de la inferencia Bayesiana presentada es general y puede ser extendida a otros modelos lineales-bilineales fijando algunos parámetros igual a cero y relajando algunas restricciones del modelo.

Palabras clave: Inferencia Bayesiana; Términos bilineales de la interacción; Tabla de doble entrada con interacción; von Mises-Fisher.

Capítulo publicado como artículo científico en JABES doi: 10.1007/s13253-011-0063-9 (2011).

ABSTRAC

Statistical analyses of two-way tables with interaction arise in many different fields of research. This study proposes the von Mises-Fisher distribution as a prior on the set of orthogonal matrices in a linear-bilinear model for studying and interpreting interaction in a two-way table. Simulated and empirical plant breeding data were used for illustration; the empirical data consist of a multi-environment trial established in two consecutive years. For the simulated data, vague but proper prior distributions were used, and for the real plant breeding data, observations from the first year were used to elicit a prior for parameters of the model for data of the second year trial. Bivariate Highest Posterior Density (HPD) regions for the posterior scores are shown in the biplots and the significance of the bilinear terms was tested using the Bayes factor. Results of the plant breeding trials show the usefulness of this general Bayesian approach for breeding trials and for detecting groups of genotypes and environments that cause significant genotype × environment interaction. The present Bayes inference methodology is general and may be extended to other linear-bilinear models by fixing certain parameters equal to zero and relaxing some model constraints.

Key words: Bayesian inference; Bilinear interaction terms, Two-way table with interaction; von-Mises-Fisher.

1. INTRODUCTION

Statistical analyses of two-way tables with interactions are performed in different areas of research, for example, in agriculture, plant breeding and genetics, medicine, and the social sciences. Models combining linear and bilinear terms have proved to be useful for analyzing two-factor studies with interaction, especially when the row and column factors do not have specific structures that might suggest contrasts between rows and columns or response functions (Cornelius and Seyedsadr, 1997). This is particularly important in plant breeding, where genotypes (rows) are evaluated in several environments (columns), and genotype × environment interactions (GE) usually complicate selection decisions for the next cycle of improvement.

The usual two-way analysis of variance model is

$$\bar{y}_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \bar{\varepsilon}_{ij} \qquad (1.1)$$

where μ, α_i, β_j, and $(\alpha\beta)_{ij}$ (for i=1,2,…,r; and j=1,2,…,c) are the grand mean, the effect of the ith row, the effect of the jth column, and the effect of the interaction of the ith row on the jth column, respectively. The residuals $\bar{\varepsilon}_{ijl}$ are identically and independently distributed with $N(0, \sigma_e^2 / n_{ij})$ (for simplicity in what follows, we assumed an equal number of observations n in each cell). Parsimonious modeling of the interaction can be considered by the singular value decomposition of $(\alpha\beta)_{ij}$ and by retaining only the first few components. This gives rise to the usual linear (additive) bilinear (non-additive) two-way model originally introduced by Gollob (1968) and Mandel (1969, 1971) and extensively used in plant breeding trials for assessing adaptation and stability (Kempton 1984; Gauch 1988; Crossa Yang and Cornelius 2004; Cornelius, Crossa and Seyedsadr 1996). This is known as the Additive Main effect and Multiplicative Interaction Model (AMMI)

$$\bar{y}_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk} + \bar{\varepsilon}_{ij} \qquad (1.2)$$

where λ_k is the singular value for the k principal component axis subject to $\lambda_1 \geq ... \lambda_t \geq 0$; u_{ik} and v_{jk} are elements of the left and right singular vectors, respectively, with the side condition that

$\sum_i u_{ik}^2 = \sum_j v_{jk}^2 = 1$ and, for $k \neq k'$, $\sum_i \alpha_{ik}\alpha_{ik'} = \sum_j \beta_{jk}\beta_{jk'} = 0$ and $t = min(r,c) - 1$. In matrix notation, (2) can be expressed as

$$\mathbf{Y} = \mu\mathbf{1}_r\mathbf{1}_c' + \boldsymbol{\alpha}\otimes\mathbf{1}_c' + \boldsymbol{\beta}'\otimes\mathbf{1}_r + \mathbf{UDV}' + \mathbf{E} \qquad (1.3)$$

where $\mathbf{Y} = [\bar{y}_{ij}]$, $\boldsymbol{\alpha} = [\alpha_i]$, $\boldsymbol{\beta} = [\beta_j]$, $\mathbf{D} = \text{diag}(\lambda_k, k = 1,2,...,t)$, $\mathbf{U} = (\boldsymbol{u}_1,...,\boldsymbol{u}_t)$, $\boldsymbol{u}_k = [u_{ik}]$, $\mathbf{V} = (\boldsymbol{v}_1,...,\boldsymbol{v}_t)$, $\boldsymbol{v}_k = [v_{jk}]$, and $\mathbf{E} = [\bar{\varepsilon}_{ij}]$. Note that from model (1.3) [or (1.2)] and allowing $\boldsymbol{\alpha} = \mathbf{0}$ and $\boldsymbol{\beta} = \mathbf{0}$, other linear-bilinear models can be obtained. For example, dropping α_i in (1.2) and writing $\mu_j = \mu + \beta_j$ gives the column regression model that has proved to be very useful in plant breeding for assessing the stability of genotypes (rows) when tested under different environmental conditions (columns). Similarly, dropping β_j in (1.2) and writing $\mu_i = \mu + \alpha_i$ gives the row regression model, and dropping μ, α_i and β_j gives the complete multiplicative model.

Commonly, the parameters in (1.3) are estimated by an iterative least squares (LS) method that first fits the linear terms while ignoring the bilinear terms, which are subsequently fitted as the first *t* components of the singular value decomposition of the residual matrix $\mathbf{Z} = \mathbf{Y} - \hat{\mu}\mathbf{1}_r\mathbf{1}_c' - \hat{\boldsymbol{\alpha}}\otimes\mathbf{1}_c' - \hat{\boldsymbol{\beta}}'\otimes\mathbf{1}_r$, where $\hat{\mu}$, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are the LS estimates obtained in the first step (Gabriel 1978). Interestingly, Seyedsadr and Cornelius (1992) showed the LS estimates of the model $\bar{y}_{ij} = \gamma + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk} + \bar{\varepsilon}_{ij}$, which was originally a problem unsolved by Gabriel (1978) (and named it the Shifted Multiplicative Model, SHMM). The main difficulties of the standard frequentist fixed or mixed linear-bilinear model (1.3) and other related models are: insufficient flexibility to handle heterogeneity of variances and unequal cell size; incorporation of previous information is not possible; only approximate tests for determining the number of components to be retained in the model are available; and inferential statistics to the interaction parameters λ_k, u_{ik}, and v_{jk} not easily developed.

Viele and Srinivasan (2000) were the first to propose Bayesian estimation of parameters for model (1.3) using MCMC techniques through Gibbs sampling with embedded Metropolis-

Hastings random walks. The authors proposed spherical uniform prior distributions for the bilinear effects and used the posterior means as shrinkage estimates. Some practical and theoretical issues unresolved by Viele and Srinivasan (2000), such as whether the MCMC will always converge on the target posterior distribution, or whether the bilinear terms of model (1.3) can be estimated from the MCMC sample without violating model constraints, were investigated in the unpublished Ph.D. thesis of Liu (2001). Liu used the same prior distributions as Viele and Srinivasan (2000) to derive the posterior conditional distributions for model (1.3) parameters, such that a Gibbs sampler (without using embedded Metropolis-Hastings steps) for sampling from the joint posterior distribution could be used. Recently, Crossa et al. (2011) applied this approach to practical data resulting from plant breeding multi-environment trials and showed that inferential statistics can be incorporated naturally by adopting the Bayesian approach for estimating the GE interaction parameters including confidence regions in the biplot of the first two components of the bilinear terms.

The spherical uniform distribution is a special case of the von Mises-Fisher distribution (Mardia, Kent, and Bibbi 1979) and was used as a prior by Viele and Srinivasan (2000) for estimating the interaction parameters of (1.3). The authors referred to the constraints on the vectors $\boldsymbol{u}_k = [u_{ik}]$ and $\boldsymbol{v}_k = [v_{jk}]$, which must have unit length and zero sum; in other words, $\boldsymbol{u}_k = [u_{ik}]$ and $\boldsymbol{v}_k = [v_{jk}]$ must be orthonormal and orthogonal to the **1** vector. However, the support of the joint posterior distribution of $\boldsymbol{u}_k = [u_{ik}]$ and $\boldsymbol{v}_k = [v_{jk}]$ is not trivial, and Viele and Srinivasan (2000) described a solution for sampling from it using the correct supports. Conceptually, this approach to sampling the conditional posterior distributions of $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$, which are spherical distributions, while maintaining the constraints on the parameters, was performed within the vector framework, that is, sequentially for column vectors of **U** and **V**.

However, there is a need to consider probability models for data from higher dimensionality that allow generalization of the vector approach to a matrix approach; specifically, a useful method would be to use a von Mises-Fisher distribution as a prior on the set of orthonormal matrices whose terms are the bilinear coefficients. Hoff (2009) showed how to sample from the von Mises-Fisher distribution on the multi-dimensional sphere considering the posterior distributions of orthonormal matrices that arise in the analysis of multivariate data.

In this paper we show how to adopt and use the multivariate von Mises-Fisher distribution as a prior on the set of orthonormal matrices that produce the singular value decomposition of interaction matrices as those suitable for model (1.3). In Section 2 we define the joint prior for unknowns in model (1.3) using the multivariate von Mises-Fisher distribution as a conditional prior for the bilinear effects, the conditional posterior distributions, the Gibbs sampling scheme to simulate MCMC samples from the joint posterior distribution, and the Bayes factor for testing the number of bilinear terms to be retained in the model. Section 3 shows the results and discussion of the application in the context of one simulated data set (five rows and three columns with interaction) and one plant breeding multi-environment multi-year trial comprising 12 genotypes and 25 environments evaluated in two consecutive years, where data from the first year are formally incorporated into the Bayesian inference process through the prior distribution when analyzing the new data from the second year. Bivariate confidence regions (HPD) were estimated for the first two components of the GE bilinear interaction parameters, and only those with HPD not covering the origin (0, 0) are shown in graphical form (biplot). Extension to other type of linear-bilinear models is discussed. Section 4 gives the conclusions.

2. BAYESIAN INFERENCE FOR THE AMMI MODEL

2.1 LIKELIHOOD FUNCTION

The likelihood function for parameters of model (1.3) is

$$\mathrm{L}(\mu,\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{U},\mathbf{V},\mathbf{D},\tau,|\,\mathbf{Y})\propto\tau^{\frac{nrc}{2}}\exp\left\{-\frac{\tau}{2}\left[n\mathrm{tr}(\mathbf{E}\mathbf{E}')+(n-1)\mathrm{tr}(\mathbf{S}\mathbf{S}')\right]\right\} \qquad (2.1)$$

where $\tau=1/\sigma^2$, $\mathbf{S}=\left\{\sqrt{s_{ij}^2}\right\}$, $\mathrm{s}_{\mathrm{ij}}^2=\sum_{l=1}^{n}\frac{(\bar{\mathbf{Y}}_{ij}-Y_{ijl})^2}{n-1}$ and $\mathbf{E}=\mathbf{Y}-\mu\mathbf{1}_r\mathbf{1}_c'-\boldsymbol{\alpha}\otimes\mathbf{1}_c'-\boldsymbol{\beta}'\otimes\mathbf{1}_r-\mathbf{U}\mathbf{D}\mathbf{V}'$.

2.2 PRIOR DISTRIBUTION

For assessing the prior distributions of the unknowns, we used conditional conjugate prior distributions such that the posterior distribution is proper and can be used to incorporate valuable

prior information from experimenters' expertise or from information generated by previous trials. Note that, since the matrices **U** and **V** are orthonormal and **D** is diagonal,

$$\begin{aligned}
&\mathrm{tr}(\mathbf{1}_r\mathbf{1}_c'\mathbf{VDU}') = 0\\
&\mathrm{tr}((\boldsymbol{\alpha}\otimes\mathbf{1}_c')\mathbf{VDU}') = 0\\
&\mathrm{tr}((\boldsymbol{\beta}'\otimes\mathbf{1}_r)\mathbf{VDU}') = 0\\
&\mathrm{tr}((\mathbf{UDV}')(\mathbf{UDV}')') = \mathrm{tr}(\mathbf{D}'\mathbf{D}) = \sum_{k=1}^{t}\lambda_k^2\\
&\mathrm{tr}\{(-2\mathbf{Y}+\mathbf{UDV}')(\mathbf{UDV}')'\} = \mathrm{tr}\{(\mathbf{D}-\mathbf{U}'\mathbf{YV})'(\mathbf{D}-\mathbf{U}'\mathbf{YV}) - (\mathbf{U}'\mathbf{YV})'(\mathbf{U}'\mathbf{YV})\}
\end{aligned} \tag{2.2}$$

thus it can be shown from (1.3) that, given $\boldsymbol{\theta} = (\mu, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and τ, the conditional likelihood function for the matrices (**U**, **D**, **V**) is

$$\begin{aligned}
L(\mathbf{U},\mathbf{V},\mathbf{D}\,|\,\boldsymbol{\theta},\tau,\mathbf{Y}) = L(\mathbf{U},\mathbf{V},\mathbf{D}\,|\,\tau,\mathbf{Y}) &\propto \exp\{-\tfrac{n\tau}{2}\,\mathrm{tr}((-2\mathbf{Y}+\mathbf{UDV}')(\mathbf{UDV}')')\}\\
&= \mathrm{etr}\{-\tfrac{n\tau}{2}(-2\mathbf{Y}+\mathbf{UDV}')(\mathbf{UDV}')'\}
\end{aligned} \tag{2.3}$$

where 'etr' is the exponential of the trace. From (5) and (6), the conditional likelihoods for $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{D}$ are

$$L(\mathbf{U}\,|\,\mathbf{V},\mathbf{D},\tau,\mathbf{Y}) \propto \mathrm{etr}\{n\tau\mathbf{YVDU}'\} \tag{2.4}$$

$$L(\mathbf{V}\,|\,\mathbf{U},\mathbf{D},\tau,\mathbf{Y}) \propto \mathrm{etr}\{n\tau\mathbf{Y}'\mathbf{UDV}'\} \tag{2.5}$$

$$\begin{aligned}
L(\mathbf{D}\,|\,\mathbf{U},\mathbf{V},\tau,\mathbf{Y}) &\propto \mathrm{etr}\{-\tfrac{n\tau}{2}(\mathbf{D}-\mathbf{U}'\mathbf{YV})'(\mathbf{D}-\mathbf{U}'\mathbf{YV})\}\\
&\propto \exp\{-\tfrac{n\tau}{2}\sum_{k=1}^{t}(\lambda_k - l_k)^2\},\ \lambda_1 > \lambda_2 > \ldots > \lambda_t,
\end{aligned} \tag{2.6}$$

respectively, where $(l_1, \ldots, l_t) = diag(\,\mathbf{U}'\mathbf{YV}\,)$.

From expression (2.4) it follows that a conditional conjugate prior for $\mathbf{U}$ is

$$\pi(\mathbf{U}\,|\,\tau) \propto \mathrm{etr}\{\tau n_0\mathbf{Y}_0\mathbf{V}_0\mathbf{D}_0\mathbf{U}'\}. \tag{2.7}$$

where $\mathbf{Y}_0$ could be interpreted as the matrix of prior cell averages such that

$$\mathbf{Y}_0 = \mu_0\mathbf{1}_r\mathbf{1}_c' + \boldsymbol{\alpha}_0\otimes\mathbf{1}_c' + \boldsymbol{\beta}_0'\otimes\mathbf{1}_r + \mathbf{U}_0\mathbf{D}_0\mathbf{V}_0';$$

that is, $\mathbf{D}_0$ is a diagonal matrix of prior singular values, and $\mathbf{U}_0$ and $\mathbf{V}_0$ are orthonormal matrices such that $\mathbf{U}_0\mathbf{D}_0\mathbf{V}_0'$ is the SVD decomposition of $\mathbf{Z}_0 = \mathbf{Y}_0 - \mu_0\mathbf{1}_r\mathbf{1}_c' - \boldsymbol{\alpha}_0 \otimes \mathbf{1}_c' - \boldsymbol{\beta}_0' \otimes \mathbf{1}_r$, where μ_0, $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ are prior values for the linear effects.

Similarly, from (2.5), a prior for $\mathbf{V}$ is

$$\pi(\mathbf{V} \mid \tau) \propto \text{etr}\{\tau n_0 \mathbf{Y}_0'\mathbf{U}_0\mathbf{D}_0\mathbf{V}'\}. \tag{2.8}$$

Both (2.7) and (2.8) are von Mises-Fisher distributions (see Appendix A). From (2.6) for each one of the elements $\lambda_1 > \lambda_2 > ... > \lambda_k$ on the diagonal of $\mathbf{D}$, the conditional conjugate prior distributions are left truncated normal with marginal densities of the form

$$\pi(\lambda_k \mid \tau) = \left\{1 - \Phi\left(\sqrt{n_0\tau}(\lambda_{k+1} - l_k^0)\right)\right\}^{-1} \mathrm{N}\left(\lambda_k \mid l_k^0, (n_0\tau)^{-1}\right),\ k = 1, ..., t,\ \lambda_{t+1} = 0 \tag{2.9}$$

For the linear terms $\boldsymbol{\theta} = (\mu, \boldsymbol{\alpha}, \boldsymbol{\beta})$ of model (1.1), a conditional conjugate prior is a $(1 + r + c)$-multivariate normal distribution with mean $\boldsymbol{\theta}_0 = (\mu_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0)$ and singular block diagonal covariance matrix

$$(n_0\tau)^{-1}\begin{bmatrix} (r_0c_0)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & c_0^{-1}\mathbf{K}_r\mathbf{K}_r^{'} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & r_0^{-1}\mathbf{K}_c\mathbf{K}_c^{'} \end{bmatrix},$$

where $\mathbf{K}_w$ is a matrix such that $\mathbf{K}_w'\mathbf{K}_w = \mathbf{I}_{w-1}$ and $\mathbf{K}_w\mathbf{K}_w' = \mathbf{I}_w - \frac{1}{w}\mathbf{J}_w$, where $\mathbf{J}_w$ is an $w \times w$ matrix with all its elements equal to one. Because of the restrictions $\boldsymbol{\alpha}'\mathbf{1}_r = 0$ and $\boldsymbol{\beta}'\mathbf{1}_c = 0$, the distribution characterized by the covariance matrix above is a singular multivariate normal distribution that does not have a density. For a prior density we need to choose a one to one transformation like $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) = (\mathbf{K}_r'\boldsymbol{\alpha}, \mathbf{K}_c'\boldsymbol{\beta})$.

Let $\boldsymbol{\theta}^* = (\mu, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$; then the prior density of $\boldsymbol{\theta}^*$ is

$$\pi(\boldsymbol{\theta}^* \mid \tau) \propto |\Sigma_0|^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0^*)^t\Sigma_0^{-1}(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0^*)\} \tag{2.10}$$

$$\Sigma_0 = (n_0\tau)^{-1}\begin{bmatrix}(r_0c_0)^{-1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & c_0^{-1}\mathbf{I}_{r-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & r_0^{-1}\mathbf{I}_{c-1}\end{bmatrix}$$

which is the density of a $(1+r-1+c-1)$-multivariate normal distribution with mean $\boldsymbol{\theta}_0^* = (\mu_0, \boldsymbol{\alpha}_0^*, \boldsymbol{\beta}_0^*) = (\mu, \mathbf{K}_r'\boldsymbol{\alpha}_0, \mathbf{K}_c'\boldsymbol{\beta}_0)$ and covariance matrix Σ_0.

The joint likelihood (2.1) suggests that a conjugate prior for τ is a gamma distribution with parameters $a/2$, and $as_0^2/2$; that is,

$$\pi(\tau) \propto \tau^{\frac{a}{2}-1}\exp\{-\frac{as_0^2}{2}\tau\} \qquad (2.11)$$

or equivalently, $as_0^2\tau \sim \chi_a^2$; thus, a and s_0^2 may be considered as prior values for sample size and variance, respectively.

Finally, the joint prior for $(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau)$ is

$$\pi\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau\right) = \pi(\boldsymbol{\theta}^*|\tau)\pi(\mathbf{U}|\tau)\pi(\mathbf{D}|\tau)\pi(\mathbf{V}|\tau)\pi(\tau) \qquad (2.12)$$

The proposed prior has practical advantages and is flexible enough to incorporate prior uncertainty about unknown parameters. On the other hand, the main disadvantage of the prior used by Liu (2001) and Crossa et al. (2011) for implementing their Bayesian approach was the elicitation of the distribution of each element on the matrices given by the SVD decomposition of the interaction. In our proposal, the incorporation of prior information is straightforward and intuitive, as it only needs to express our beliefs in the prior cell averages $\mathbf{Y}_0$ and prior linear effects $\boldsymbol{\theta}_0$. Then $\mathbf{U}_0$, $\mathbf{V}_0$ and $\mathbf{D}_0$ are obtained from the SVD decomposition of $\mathbf{Z}_0$, under the restrictions $\mathbf{U}_0'\mathbf{1}_r = \mathbf{0}$ and $\mathbf{V}_0'\mathbf{1}_c = \mathbf{0}$. The prior distribution of the linear effects is completely specified by giving a belief $\boldsymbol{\theta}_0^*$ about $\boldsymbol{\theta}^*$. This prior belief may be expressed as a function of $\mathbf{Y}_0$; for example, $\boldsymbol{\theta}_0^*(\mathbf{Y}_0) = \left(\frac{\mathbf{1}_r'\mathbf{Y}_0\mathbf{1}_c}{r_0c_0}, \frac{\mathbf{K}_r'\mathbf{Y}_0\mathbf{1}_c}{c_0}, \frac{\mathbf{K}_c'\mathbf{Y}_0'\mathbf{1}_r}{r_0}\right)$. Also, it is important to note that a vague prior for τ implies diffuse priors for all the other parameters. Then, an objective or default Bayesian analysis could be performed by setting small values to the hyper-parameter a in the prior for τ given by (2.11). Therefore, we may summarize our prior information by giving, a priori, a

prediction of the two-way array of means $\mathbf{Y}_0$ and a measure of our prior uncertainty s_0^2 given a prior sample size a.

For the analysis of the simulated data set below, the priors used were vague but proper, whereas for the plant breeding data set, the prior hyper-parameters used were obtained from the first-year evaluation of the 12 genotypes in 25 environments, and the data set analyzed was from the second year of evaluation. This example illustrates how to use this approach within a practical breeding program and how to draw useful biological inferences on the interaction parameters.

2.3 POSTERIOR DISTRIBUTION AND GIBBS SAMPLER

The joint posterior distribution is obtained by combining the likelihood function (2.1) and the prior distribution (2.12),

$$\pi\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right) \propto L\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right) \pi(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau)$$

where $L\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right)$ is a re-parameterization of $L\left(\boldsymbol{\theta}, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right)$. The marginal posterior distribution, which involves high dimensional integration on complex spaces, is needed for marginal inference about the unknowns. In order to use a Markov Chain Monte Carlo (MCMC) method through the Gibbs sampler, the full conditional posterior distributions, which are known except for the proportionality constants, are needed. These distributions are computed by considering the joint posterior as a function only of a variable when fixing the others. Thus, the conditional posterior for $\boldsymbol{\theta}^*$ is

$$\pi(\boldsymbol{\theta}^* | \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau, \mathbf{Y}) \propto L\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right) \pi\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau\right)$$

Knowing that the conditional likelihood function of $(\mathbf{U}, \mathbf{V}, \mathbf{D})$ does not depend on $\boldsymbol{\theta}^*$**,** and that given τ, the prior for $\boldsymbol{\theta}^*$ is independent of $(\mathbf{U}, \mathbf{V}, \mathbf{D})$, then

$$\pi(\boldsymbol{\theta}^* | \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau, \mathbf{Y}) = \pi(\boldsymbol{\theta}^* | \tau, \mathbf{Y}) \propto L(\boldsymbol{\theta}^* | \tau, \mathbf{Y}) \pi\left(\boldsymbol{\theta}^* \middle| \tau\right)$$

$$\propto L\left(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right) \pi\left(\boldsymbol{\theta}^* \middle| \tau\right) / L(\mathbf{U}, \mathbf{V}, \mathbf{D} | \tau, \mathbf{Y})$$

$$\propto \exp\{-\frac{n\tau}{2}\mathrm{tr}(\mathbf{ZZ}')\}\,\pi\left(\boldsymbol{\theta}^{*}\,\middle|\tau\right),$$

where $\mathbf{Z}=\mathbf{Y}-\mu\mathbf{1}_r\mathbf{1}_c'-\boldsymbol{\alpha}\otimes\mathbf{1}_c'-\boldsymbol{\beta}'\otimes\mathbf{1}_r\Big|_{\boldsymbol{\alpha}=\mathbf{K}_r\boldsymbol{\alpha}^{*},\boldsymbol{\beta}=\mathbf{K}_c\boldsymbol{\beta}^{*}}$. It can be shown that the conditional posterior of $\boldsymbol{\theta}^{*}$ is multivariate normal with density $\pi(\boldsymbol{\theta}^{*}|\tau,\mathbf{Y})=N_{r+c-1}(\boldsymbol{\theta}^{*}|\boldsymbol{\theta}_n^{*},\boldsymbol{\Sigma}_n^{*})$, covariance matrix $\boldsymbol{\Sigma}_n^{*}=(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_{\boldsymbol{n}}^{-\mathbf{1}})^{-1}$ and mean $\boldsymbol{\theta}_n^{*}=(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_{\boldsymbol{n}}^{-\mathbf{1}})^{-1}\times\left(\boldsymbol{\Sigma}_{\mathbf{0}}^{-\mathbf{1}}\boldsymbol{\theta}_0^{*}+\boldsymbol{\Sigma}_{\boldsymbol{n}}^{-\mathbf{1}}\widehat{\boldsymbol{\theta}}_n^{*}\right)$ with $\widehat{\boldsymbol{\theta}}_n^{*}=\left(\frac{\mathbf{1}_r'\mathbf{Y}\mathbf{1}_c}{rc},\frac{\mathbf{K}_r'\mathbf{Y}\mathbf{1}_c}{c},\frac{\mathbf{K}_c'\mathbf{Y}'\mathbf{1}_r}{r}\right)$ and

$$\boldsymbol{\Sigma}_n=(n\tau)^{-1}\begin{bmatrix}(rc)^{-1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & c^{-1}\mathbf{I}_{r-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & r^{-1}\mathbf{I}_{c-1}\end{bmatrix}.$$

We may use the conditional likelihoods given by (2.4)-(2.6) to calculate the conditional posteriors for $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{D}$; i.e.,

$$\pi(\mathbf{U}\,|\,\mathbf{V},\mathbf{D},\tau,\mathbf{Y})\propto L(\mathbf{U}\,|\,\mathbf{V},\mathbf{D},\tau,\mathbf{Y})\pi(\mathbf{U}\,|\,\tau)\propto \mathrm{etr}\left\{\tau\left(n_0\mathbf{Y}_0\mathbf{V}_0\mathbf{D}_0+n\mathbf{YVD}\right)\mathbf{U}'\right\}, \tag{2.13}$$

$$\pi(\mathbf{V}\,|\,\mathbf{U},\mathbf{D},\tau,\mathbf{Y})\propto L(\mathbf{V}\,|\,\mathbf{U},\mathbf{D},\tau,\mathbf{Y})\pi(\mathbf{V}\,|\,\tau)\propto \mathrm{etr}\left\{\tau\left(n_0\mathbf{Y}_0'\mathbf{U}_0\mathbf{D}_0+n\mathbf{Y}'\mathbf{UD}\right)\mathbf{V}'\right\}, \tag{2.14}$$

$$\begin{aligned}\pi(\mathbf{D}\,|\,\mathbf{U},\mathbf{V},\tau,\mathbf{Y})&\propto L(\mathbf{D}\,|\,\mathbf{U},\mathbf{V},\tau,\mathbf{Y})\pi(\mathbf{D}\,|\,\mathbf{U},\mathbf{V},\tau)\propto \exp\left\{-\tfrac{n\tau}{2}\sum_{k=1}^{t}(\lambda_k-l_k)^2\right\}\prod_{k=1}^{t}\pi(\lambda_k\,|\,\tau)\\ &\propto\prod_{k=1}^{t}\left\{1-\Phi\left(\sqrt{\frac{\tau^{-1}}{n_0+n}}\left(\lambda_{k+1}-\frac{n_0l_k^0+nl_k}{n_0+n}\right)\right)\right\}^{-1}\mathrm{N}\left(\lambda_k\,|\,\frac{n_0l_k^0+nl_k}{n_0+n},\frac{\tau^{-1}}{n_0+n}\right),\\ &\qquad\qquad \lambda_1>\lambda_2>\ldots>\lambda_t>\lambda_{t+1}=0\end{aligned} \tag{2.15}$$

Finally, the conditional posterior for the precision τ is a gamma with density

$$\begin{aligned}\pi(\tau\,|\,\boldsymbol{\theta}^{*},\mathbf{U},\mathbf{V},\mathbf{D},\mathbf{Y})&\propto \mathrm{L}(\boldsymbol{\theta}^{*},\mathbf{U},\mathbf{V},\mathbf{D},\tau\,|\,\mathbf{Y})\pi(\tau)\\ &\propto \mathrm{Ga}\left(\tau|\frac{a_n}{2},\frac{b_n}{2}\right),\end{aligned}$$

where

$$\begin{aligned}a_n&=a+nrc\\ b_n&=bs_0^2+n\left(\mathrm{tr}(\mathbf{EE}')\right)+(n-1)\mathrm{tr}(\mathbf{SS}')\end{aligned}$$

2.3.1 Gibbs sampler

The Gibbs Sampler is implemented by sequentially drawing simulated samples from the full conditional posterior distributions; thus we may proceed with the following algorithm:

Let *s* be the desired length of the Markov chain to be simulated. Let $\mathbf{U}^{(0)}$, $\mathbf{V}^{(0)}$ and $\mathbf{D}^{(0)}$ be the initial values of the simulated Markov chain.

For $i=0,\ldots,s$ simulate

$$\tau^{(i+1)} \sim \pi(\tau \mid \boldsymbol{\theta}^{*(i)}, \mathbf{U}^{(i)}, \mathbf{V}^{(i)}, \mathbf{D}^{(i)}, \mathbf{Y})$$

$$\boldsymbol{\theta}^{*(i+1)} \sim \pi\left(\boldsymbol{\theta}^{*} \middle| \tau^{(i+1)}, \mathbf{Y}\right)$$

$$\mathbf{U}^{(i+1)} \sim \pi(\mathbf{U} \mid \tau^{(i+1)})$$

$$\mathbf{D}^{(i+1)} \sim \pi(\mathbf{D} \mid \tau^{(i+1)})$$

$$\mathbf{V}^{(i+1)} \sim \pi(\mathbf{V} \mid \tau^{(i+1)})$$

After a burn-in period, we assume that the generated samples arise from the stationary distribution, i.e., the joint posterior distribution $\pi\left(\boldsymbol{\theta}^{*}, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tau \middle| \mathbf{Y}\right)$. For reversibility of the Markov chain, we may permute the order of simulation, but in what follows we use the order indicated above. Some standard convergence diagnostic tools may be used to determine an effective sample size; in the examples below, criteria for convergence of simulated Markov chains from Raftery and Lewis (1995) and Gelman and Rubin (1992) were used.

2.4 TESTING THE SIGNIFICANCE OF THE BILINEAR TERMS USING THE BAYES FACTOR

Concerning the inference on the parameters of liner-bilinear models, a central question is how many linear components the model should retain. Several frequentist model selection approaches exist for linear-bilinear models of a two-way table without replicated data in each cell. Johnson and Graybill (1972) developed a hypothesis test for a model with one bilinear term

and Marasinghe (1985) and Schott (1986) proposed a sequential test for the bilinear terms beyond the first one. For two-way tables with replications in each cell, Gollob (1968) proposed using an approximate *F* statistics for the hypothesis H_1: $\lambda_t = 0$ versus H_2: $\lambda_t \neq 0$. However, this test is too liberal for practical use and Cornelius, Seyedsadr, and Crossa (1992), and Cornelius, Crossa, and Seyedsadr (1994) introduced a series of sequential *F* approximations that effectively control Type I error rates.

Several approaches can be found in the literature for Bayesian model selection, being the most widely used the based on the Bayes factors. Suppose a data sample (***Y***) coming from two competing models M_0 and M_1 according to probabilities $p(\boldsymbol{Y}|M_0)$ and $p(\boldsymbol{Y}|M_1)$, whose prior probabilities are such that $p(M_0)$ and $p(M_1)$=1- $p(M_0)$, and their respective posterior probabilities are $p(M_0 \,|\boldsymbol{Y})$ and $p(M_1|\ \boldsymbol{Y})$. The Bayes factor B_{01}, defined as the ratio of the posterior to prior odds, is the weight of evidence in favor of model M_1 provided by the data (Kass and Raftery, 1995); that is,

$$B_{01} = \frac{p(M_0 \,|\, \mathbf{Y})}{p(M_1 \,|\, \mathbf{Y})} \Big/ \frac{p(M_0)}{p(M_1)} = \frac{p(\mathbf{Y} \,|\, M_0)}{p(\mathbf{Y} \,|\, M_1)}.$$

The conventional decision rule suggested by Jeffrerys (1961) is that M_0 is selected if B_{10}>10 otherwise, the model selected is M_1. Note that this reduces to the posterior odds in favor of M_0 when M_0 and M_1 have equal prior probabilities, $p(M_0)$=$p(M_1)$=0.5 (representing the usual noninformative prior on two competitive models). Considering that $\lambda_1 > \lambda_2 \ldots > \lambda_t$, it is sufficient to sequentially compare model M_0 with $\lambda_k \neq 0$ and $\lambda_{k+1} = 0$ against the alternative model M_1 with $\lambda_{k+1} \neq 0$ [for *k*=1,2,…,*t*] in order to choose the model with highest posterior probability. By the above definition of Bayes factor, the interest lies in obtaining the marginal density of the data $p(\mathbf{Y} \,|\, M_i)$ for which Chib (1995) provided a method for computing this when the full conditionals of blocks of parameters are available in closed form. For a model where **Y** is the observed data and **θ** denotes the unknown parameter, Chib (1995) noted that by the Bayes's rule "**θ**Î Q $p(\mathbf{Y} \,|\, M_i) = f(\mathbf{Y} \,|\, \boldsymbol{\theta}, M_i)\, p(\boldsymbol{\theta} \,|\, M_i) / p(\boldsymbol{\theta} \,|\, \mathbf{Y}, M_i)$. Then, an estimate of log $p(\mathbf{Y} \,|\, M_i)$ is

$$\log p(\mathbf{Y} \mid M_i) = \log f(\mathbf{Y} \mid \boldsymbol{\theta}, M_i) + \log p(\boldsymbol{\theta} \mid M_i) - \log p(\boldsymbol{\theta} \mid \mathbf{Y}, M_i)$$

where $p(\boldsymbol{\theta} \mid \mathbf{Y}, M_i)$ is an estimate of the posterior distribution under the model M_i and $\boldsymbol{\theta}$ is a value of $\boldsymbol{\theta}$ with high posterior density to assure accuracy. It should be clear that normalization constants need to be known or at least estimated for each model. Therefore, an approximation of B_{01} is given by

$$B_{01} = \frac{p(\mathbf{Y} \mid M_0)}{p(\mathbf{Y} \mid M_1)}$$

For a detailed explanation of the algorithm and alternative approaches for computing Bayes factors, see Han and Carlin (2001). Since this algorithm works with the models one at a time, model indicators are dropped to simplify the notation. Note that in our problem

$$p(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, t \mid \mathbf{Y}) = p(\mathbf{D} \mid \mathbf{U}, \mathbf{V}, \boldsymbol{\theta}^*, t, \mathbf{Y}) p(\mathbf{U} \mid \mathbf{V}, \boldsymbol{\theta}^*, t, \mathbf{Y}) p(\mathbf{V} \mid \boldsymbol{\theta}, t, \mathbf{Y}) p(\boldsymbol{\theta}^* \mid t, \mathbf{Y}) p(t \mid \mathbf{Y})$$

Thus, an estimate of the joint posterior at the high density point $(\hat{\boldsymbol{\theta}}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{D}}, \hat{t})$ is

$$\begin{aligned} \hat{p}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\mathbf{D}}, \hat{t} \mid \mathbf{Y}) &= p(\hat{\mathbf{D}} \mid \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{\boldsymbol{\theta}}, \hat{t}, \mathbf{Y}) \hat{p}(\hat{\mathbf{U}} \mid \hat{\mathbf{V}}, \hat{\boldsymbol{\theta}}, \hat{t}, \mathbf{Y}) \hat{p}(\hat{\mathbf{V}} \mid \hat{\boldsymbol{\theta}}, \hat{t}, \mathbf{Y}) p(\hat{\boldsymbol{\theta}} \mid \hat{t}, \mathbf{Y}) \hat{p}(\hat{t} \mid \mathbf{Y}) \\ &= p(\hat{\mathbf{D}} \mid \hat{\mathbf{U}}, \hat{\mathbf{V}}, \hat{t}, \mathbf{Y}) \hat{p}(\hat{\mathbf{U}} \mid \hat{\mathbf{V}}, \hat{t}, \mathbf{Y}) \hat{p}(\hat{\mathbf{V}} \mid \hat{t}, \mathbf{Y}) p(\hat{\boldsymbol{\theta}} \mid \hat{t}, \mathbf{Y}) \hat{p}(\hat{t} \mid \mathbf{Y}) \end{aligned}$$

were

$$\hat{p}(\hat{\mathbf{U}} \mid \hat{\mathbf{V}}, \hat{t}, \mathbf{Y}) = \sum_{g=1}^{G} p(\hat{\mathbf{U}} \mid \mathbf{D}^{(g)}, \hat{\mathbf{V}}, \hat{t}, \mathbf{Y}) / G$$

$$\hat{p}(\hat{\mathbf{V}} \mid \hat{t}, \mathbf{Y}) = \sum_{g=1}^{G} p(\hat{\mathbf{V}} \mid \mathbf{D}^{(g)}, \mathbf{U}^{(g)}, \hat{t}, \mathbf{Y}) / G$$

$$\hat{p}(\hat{t} \mid \mathbf{Y}) = \sum_{g=1}^{G} p(\hat{t} \mid \mathbf{U}^{(g)}, \mathbf{D}^{(g)}, \mathbf{V}^{(g)}, \boldsymbol{\theta}^{*(g)}, \mathbf{Y}) / G$$

are Monte Carlo estimates of the conditional densities based on a Gibbs sample of size *G*. We do not need an estimate of $p(\boldsymbol{\theta}^* \mid t, \mathbf{Y})$, since the full conditional does not depend on $(\mathbf{U}, \mathbf{V}, \mathbf{D})$. In this study the most difficult task is to estimate the normalization constant of the full conditionals of **U** and **V,** which are von Mises-Fisher distributions, because no closed form of the

normalization constant given by the hypergeometric function of a matrix argument (Herz, 1955; James, 1964); here we used the approximation given by Khatri and Mardia (1977) for either small or large eigenvalues of the matrix argument. A more accurate approximation was suggested by Koev and Edelman (2006).

With the above approximation of the joint posterior, using the priors given in Section 2.2, and adjusting them by the model size, an estimate of the marginal log-density of **Y** conditional on M_i is

$$\log p(\mathbf{Y} \mid M_i) = \log \mathrm{L}(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tilde{\tau} \mid \mathbf{Y}, M_i) + \log \pi(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tilde{\tau} \mid M_i) - \log \pi(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{V}, \mathbf{D}, \tilde{\tau} \mid \mathbf{Y}, M_i).$$

Note that by (2.1)-(2.3) and (2.12), the approximation may be written as

$$\begin{aligned}\log p(\mathbf{Y} \mid M_i) = {} & \log \mathrm{L}(\boldsymbol{\theta}^*, \tilde{\tau} \mid \mathbf{Y}) + \log \mathrm{L}(\mathbf{U}, \mathbf{V}, \mathbf{D} \mid \tilde{\tau}, \mathbf{Y}, M_i) + \log \pi(\boldsymbol{\theta}^* \mid \tilde{\tau}) + \log \pi(\mathbf{U} \mid \tilde{\tau}, M_i) \\ & + \log \pi(\mathbf{V} \mid \tilde{\tau}, M_i) + \log \pi(\mathbf{D} \mid \tilde{\tau}, M_i) + \log \pi(\tilde{\tau} \mid M_i) - \log \pi(\mathbf{D} \mid \mathbf{U}, \mathbf{V}, \tilde{\tau}, \mathbf{Y}, M_i) \\ & - \log \pi(\mathbf{U} \mid \mathbf{V}, \tilde{\tau}, \mathbf{Y}, M_i) - \log \pi(\mathbf{V} \mid \tilde{\tau}, \mathbf{Y}, M_i) - \log \pi(\boldsymbol{\theta}^* \mid \tilde{\tau}, \mathbf{Y}) - \log \pi(\tilde{\tau} \mid \mathbf{Y}, M_i).\end{aligned}$$

In the equation above the terms involving $\boldsymbol{\theta}^*$ does not depend on M_i; therefore, when estimating the Bayes factor is not necessary to evaluate that terms since it will be canceled out.

3. RESULTS AND DISCUSSION

3.1 SIMULATED DATA

To illustrate the implementation of our proposed model, we analyzed a simulated data set comprising five rows and three columns, each with an $n = 4$ sample size from a normal distribution with mean $\mu + \alpha_i + \sum_{k=1}^{2} \lambda_k u_{ik} v_{jk}$, $(i = 1, \ldots, 5; j = 1, \ldots, 3)$ and variance $\sigma^2 = 1$. Thus the overall sample size is 60. The true values of the fixed linear effects are $\alpha = (-1, -0.5, 0.5, 1, 0)$ for the factor row and $\beta = (-1, 1, 0)$ for the factor column, while the bilinear terms are given by the SVD decomposition with components:

$$\mathbf{U} = \begin{bmatrix} -0.8250 & -0.3372 \\ -0.0943 & 0.7041 \\ 0.1785 & 0.3049 \\ 0.3246 & -0.1463 \\ 0.4162 & -0.5255 \end{bmatrix},\ \mathbf{D} = \begin{bmatrix} 1.4805 & 0 \\ 0 & 0.1579 \end{bmatrix},\ \mathbf{V} = \begin{bmatrix} -0.7865 & -0.2191 \\ 0.2035 & 0.7907 \\ 0.5830 & -0.5716 \end{bmatrix}$$

The interaction plot of the simulated data shows that the row effects are not equal for all levels of the factor column, with clear evidence of interaction between row and column effects (data not shown). The true biplot of the first (Component 1) and second (Component 2) principal components, i.e., the columns of $\mathbf{PC} = \begin{pmatrix} \mathbf{U}\mathbf{D}^{1/2} \\ \mathbf{V}\mathbf{D}^{1/2} \end{pmatrix}$, shows interaction effects between rows and columns (data not shown).

We used $\alpha_0 = (0, 0,\ 0, 0, 0)$, $\beta_0 = (0,0,0)$, $a = 1$, $s_0^2 = 1000$ as prior information, and samples of the posterior distribution were drawn with the Gibbs sampler described above. Two parallel chains of size $s = 20{,}000$ were simulated using a burn-in period of size 10000. Finally, a thinning period of 1 was used. Therefore, a final MCMC sample of size 5,000 was used to estimate the posterior distribution.

Table 1. Posterior summary (Mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.90 HPD, and 0.95 HPD intervals computed with 5000 approximately independent samples from the joint posterior distribution of the simulated data (five rows and three columns).

Parameter	Mean	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	0.90 HPD interval		0.95 HPD interval	
						Lower	Upper	Lower	Upper
μ	5.0184	0.1342	4.7523	5.0175	5.2878	4.8055	5.2414	4.7584	5.2898
σ	1.1572	0.1312	0.9401	1.1461	1.4498	0.9504	1.3617	0.9224	1.4211
α_1	−1.2282	0.2713	−1.7687	−1.2268	−0.6991	−1.6437	−0.7530	−1.7421	−0.6755
α_2	−0.6733	0.2703	−1.2027	−0.6744	−0.1334	−1.1158	−0.2299	−1.2169	−0.1493
α_3	0.6239	0.2704	0.0897	0.6234	1.1536	0.1879	1.0654	0.0825	1.1429
α_4	0.8053	0.2707	0.2635	0.8078	1.3387	0.3514	1.2457	0.2674	1.3406
α_5	0.4723	0.2693	−0.0597	0.4756	0.9977	0.0283	0.9101	−0.0524	0.9990
β_1	−0.6824	0.1891	−1.0512	−0.6835	−0.3060	−0.9833	−0.3607	−1.0442	−0.3006
β_2	0.9110	0.1919	0.5323	0.9113	1.2878	0.5876	1.2178	0.5247	1.2801
β_3	−0.2286	0.1885	−0.5925	−0.2262	0.1382	−0.5313	0.0796	−0.5984	0.1299
$u_{1,1}$	−0.7251	0.1456	−0.8822	−0.7654	−0.3227	−0.8934	−0.5411	−0.8936	−0.4429
$u_{2,1}$	0.0486	0.2471	−0.4350	0.0491	0.5363	−0.3761	0.4418	−0.4145	0.5553
$u_{3,1}$	0.3568	0.2278	−0.1516	0.3785	0.7387	0.0148	0.7348	−0.0980	0.7848
$u_{4,1}$	0.0291	0.2664	−0.5018	0.0339	0.5419	−0.4052	0.4702	−0.4965	0.5463
$u_{5,1}$	0.2906	0.2329	−0.2194	0.3065	0.7075	−0.0625	0.6817	−0.1559	0.7540
$u_{1,2}$	−0.0794	0.3027	−0.6419	−0.0856	0.5312	−0.4084	0.5854	−0.6491	0.5206
$u_{2,2}$	0.4202	0.2532	0.0020	0.4207	0.8463	−0.8109	−0.0279	0.0060	0.8490
$u_{3,2}$	−0.1020	0.4324	−0.8099	−0.1219	0.6999	−0.5690	0.8001	−0.8616	0.6398
$u_{4,2}$	−0.1439	0.4835	−0.8506	−0.2008	0.7448	−0.5732	0.8802	−0.8802	0.6927
$u_{5,2}$	−0.0949	0.4480	−0.8220	−0.1154	0.7156	−0.5929	0.8072	−0.8734	0.6487
$v_{1,1}$	−0.7217	0.1321	−0.8164	−0.7698	−0.3394	−0.8165	−0.5704	−0.8165	−0.4665
$v_{2,1}$	0.4129	0.3096	−0.3190	0.4766	0.8095	−0.0152	0.8165	−0.1733	0.8165
$v_{3,1}$	0.3088	0.3162	−0.3635	0.3356	0.7980	−0.1241	0.8165	−0.2461	0.8165
$v_{1,2}$	0.0602	0.3532	−0.6508	0.0814	0.6803	−0.6089	0.5396	−0.6057	0.7196
$v_{2,2}$	0.5941	0.2177	0.0368	0.6540	0.8160	−0.8165	−0.2917	0.1700	0.8165
$v_{3,2}$	−0.6543	0.2080	−0.8163	−0.7326	−0.0962	0.3934	0.8165	−0.8165	−0.2449
λ_1	2.2079	0.4741	1.0096	2.2793	2.9221	1.5010	2.9257	1.2120	3.0237
λ_2	0.5073	0.3744	0.0132	0.4461	1.3127	0.0003	1.0357	0.0003	1.1863

Table 1 gives a summary of the marginal posterior distribution for all the parameters in the model. The true values are always within their interval estimations. Moreover, relative to the estimated standard deviations, point estimations are close enough to the true values. Thus the proposed Bayesian approach to model interaction in a two-way table gives reliable inferential answers about unknowns in the model. The second singular value is about 4 times smaller than the first singular value, and also had smaller SD than the first one. The eigenvectors for $u_{1,1}, v_{1,1}, v_{2,2}$, and $v_{3,2}$ are the only bilinear terms that do not include the null point (0, 0) for the bivariate 0.90 HPD region and thus cause most of the significant interaction. Other posterior 0.90 HPD intervals for eigenvector elements such as $u_{3,1}$ and $u_{2,2}$ do not contain zero, but the 0.90 HPD region for their corresponding scores ($u_{3,1}\sqrt{\lambda_1}, u_{3,2}\sqrt{\lambda_2}$) and ($u_{2,1}\sqrt{\lambda_1}, u_{2,2}\sqrt{\lambda_2}$) covers the null point (0, 0).

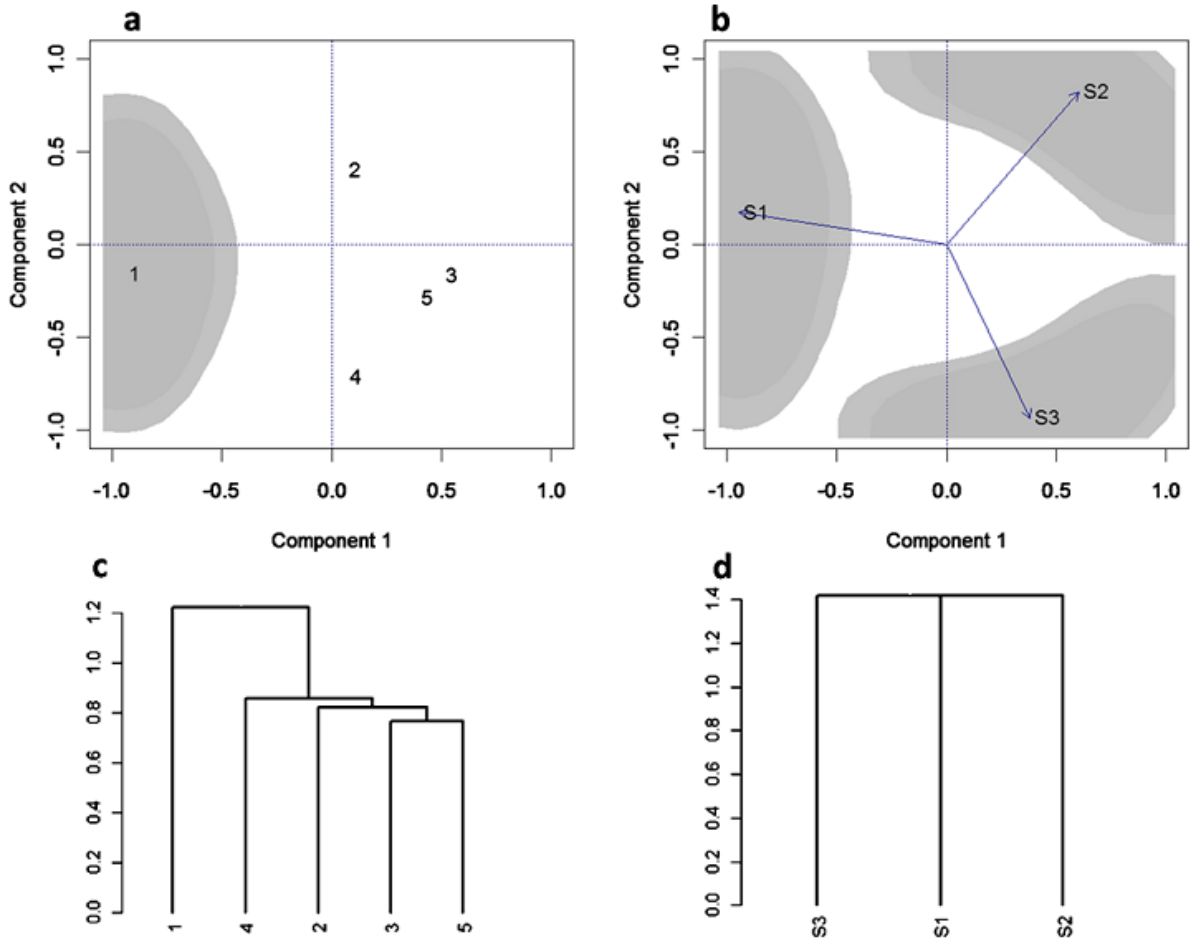


Figure 1. Simulated data with five rows (1-5) and three columns (S1-S3): (**a**) plot of the row scores $\mathbf{U}'\mathbf{D}^{1/2}$ and the bivariate 95% (gray external contour) and 90% (gray internal contour) HPD regions; (**b**) plot of the column scores $\mathbf{V}'\mathbf{D}^{1/2}$ and the bivariate .95 (gray external contour) and .90 (gray internal contour) HPD regions. Only row 1 and columns S1, S2, and S3 that do not include the null point (0, 0) at the .90 HPD probability level are depicted; (**c**) dendrogram of the five rows using the first two left singular vectors; (**d**) dendrogram of the three columns using the first two right singular vectors.

The HPD regions of the row (column) scores that are statistically different from the null point (0, 0) can be seen in the biplot depicted in Figures 1(a) and (b). For example, Figure. 1(a) shows the plot of the row scores $\mathbf{U}'\mathbf{D}^{1/2}$, and the outer and inner shaded areas of the graph are the bivariate 0.95 and 0.90 HPD posterior regions, respectively, for the scores in row 1. For clarity, the 0.90 HPD and 0.95 HPD regions for the other row scores were not drawn because they contain the null point (0, 0), which is evidence that their contribution to the interaction was not statistically significant. Analogously, from Figure 1(b) it can be seen that the posterior densities for the scores of S1, S2 and S3 do not include the null point (0, 0) at either of the two probability levels; thus we can conclude that, given the data, there is enough evidence that the multiplicative column effect is significant at all its levels. Furthermore, since there is no overlapping of the interaction scores for columns at any of the 0.90 HPD and 0.95 HPD regions, we can conclude that among these regions there are different interaction effects that are statistically significant.

As an additional descriptive tool, we performed a hierarchical cluster algorithm with a complete linkage strategy based on the Euclidean distances between the rows (and columns) of the matrix **U** (**V**); their dendrograms are presented in Figures 1(c) and (d), respectively. Row 1 is the farthest score from (0, 0) and significant for the interaction; it does not cluster with any of the other non-significant genotypes [Figure. 1(c)]. The three columns do not form any clear clusters [Figure. 1(d)].

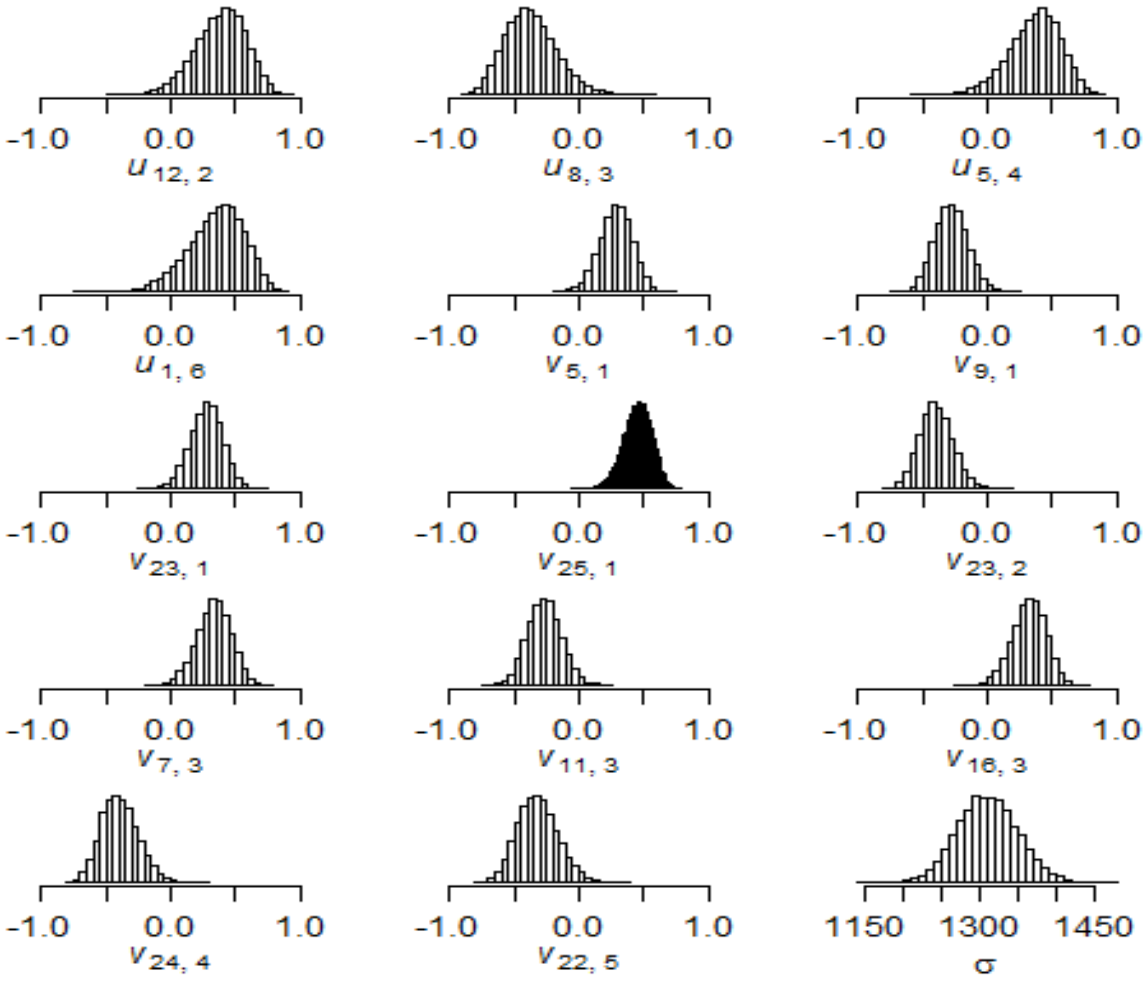


Figure 2. Plant breeding data of 12 genotypes in 25 environments. Histograms of the MCMC samples from the marginal posterior distributions of $u_{12,2}$, $u_{8,3}$, $u_{5,4}$, $u_{1,6}$, $v_{5,1}$, $v_{9,1}$, $v_{23,1}$, $v_{25,1}$, $v_{23,2}$, $v_{7,3}$, $v_{11,3}$, $v_{16,3}$, $v_{24,4}$, $v_{22,5}$ and σ.

In the next section, we illustrate the results of our prior selection and MCMC sampling strategy to analyze the genotype × environment interaction in the context of a multi-environment and multi-year plant breeding trial with genotypes evaluated under different environmental conditions during two consecutive years.

3.2 PLANT BREEDING DATA

The multi-environment multi-year plant breeding trial analyzed in this section comprises 12 maize hybrids evaluated for grain yield in 25 environments for two consecutive years. The 12 hybrids were arranged in a randomized complete block design with two replicates in each environment and year. The effect of the design (i.e. complete or incomplete block) is easily incorporated into the model.

The first-year data were used to elicit the prior to analyze data from the second year. The posterior means of the 38 linear parameters (overall mean, 12 genotypic effects and 25 environmental effects) are given in Table B.1 (Appendix B). The histograms of the posterior means of several bilinear interaction parameters depicted in Figure 2 show bell-shaped marginal posterior distributions on (-1,1). The histograms of the MCMC samples of the marginal posteriors of λ_1- λ_{11} are also shown and, as expected, the posterior densities of the late lambdas move towards zero [Figure 3(a)]. The posterior densities of the cumulative proportion of variance explained by the eigenvalues show that five components explained about 90% of the interaction variance [Figure 3(b)].

The Bayes factor computed for determining the number of λ's indicated that the model with three GE bilinear components is appropriate. Assuming that the prior odds is 1, the Bayes factor in favor of model with two components ($\lambda_2 \neq 0, \lambda_3 = 0$) when comparing against the alternative model with at least three components ($\lambda_3 \neq 0$) is 0.0022, whereas the value of the Bayes factor indicated that the model with three components ($\lambda_3 \neq 0, \lambda_4 = 0$) is 104.15 times more probable than the model with at least four components ($\lambda_4 \neq 0$). These results of three significant bilinear components ($\lambda_3 \neq 0$) is in agreement with those usually found in the analyses of plant breeding trials where the complexity of the GE requires more than one bilinear component to be retained in the model.

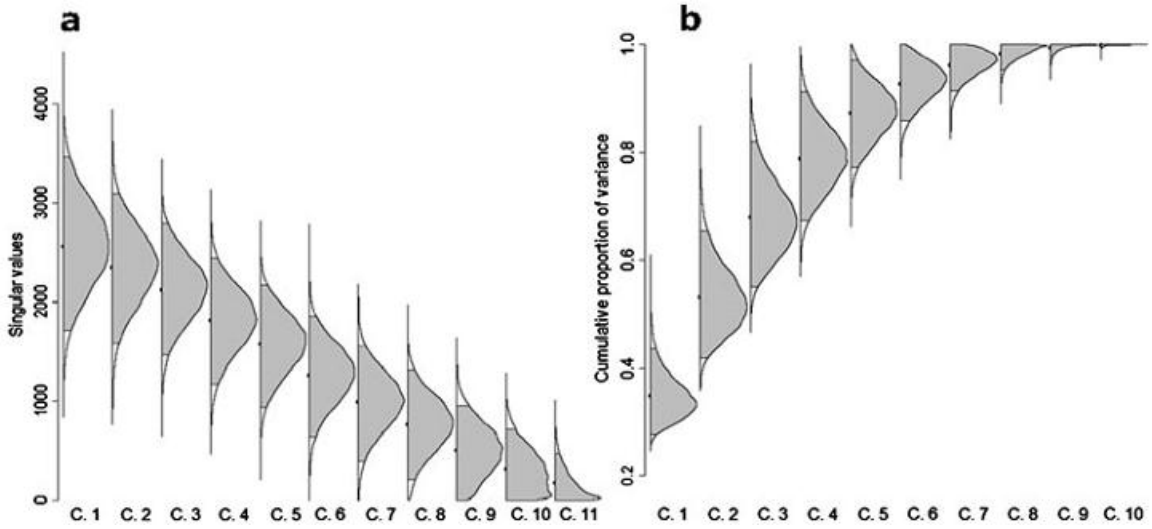


Figure 3. Plant breeding data of 12 genotypes in 25 environments: (**a**) Posterior densities and 0.95 HPD regions of the singular values, $\lambda_1, \ldots, \lambda_{11}$ (C1-C11); (**b**) Posterior densities and 0.95 HPD regions of the cumulative proportion of variance $\phi_t = \frac{\sum_{k=1}^{t} \lambda_k^2}{\sum_{k=1}^{\min(r,c)-1} \lambda_k^2}, t = 1, \ldots, \min(r,c) - 2$. The x-axes show the posterior means of the components (C1-C11) (**a**) and the cumulative components (**b**).

In general, the values of u_{i1} and v_{j1} were, in absolute terms, larger than the values of u_{i2} and v_{j2}, whereas the standard deviations (SD) of u_{i1} and v_{j1} were smaller than those of u_{i2} and v_{j2}. Consequently, the lengths of the HPD regions were narrower for the first bilinear components of genotypes and environments (u_{i1} and v_{j1}) than for the second bilinear components, u_{i2} and v_{j2} (data not shown). In summary, there is more posterior uncertainty in the second bilinear components for genotypes and sites than in the first bilinear components.

3.2.1 Credible Regions of the First Two Bilinear Terms of the Linear-Bilinear Model

Given in Table 2 are the posterior means of the 11 singular values, together with the values of the eigenvectors of genotypes and environments whose 0.95 HPD for their

corresponding scores do not contain the null value. As the singular values decrease in size, their SDs also decrease. The scores of genotype 12 ($u_{12,2}$) and the scores of environments S5, S9, S23, and S25 ($v_{5,1}, v_{9,1,}v_{23,1}, v_{25,1},$ and $v_{23,2}$) significantly contributed to genotype × environment variability, as shown by the bivariate 0.90 HPD regions, whereas the other genotypes and environments did not significantly contribute to that variability.

Posterior modes of the first and second components for the scores of environments and genotypes, with their associated HPD regions, indicate the scores that significantly contributed to the interaction between genotypes and environments (Figure 4). In this biplot, the bivariate 0.90 HPD and 0.95 HPD regions are shown only for genotype 12 [Figure 4(a)] and environments S5, S9, S23, and S25 [Figure 4(b)]. These are the genotypes and environments that contributed significantly to the interaction, as their HPD regions did not include the null point (0, 0).

Table 2. Posterior summary (Mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.90 HPD, and 0.95 HPD intervals computed with 20,000 approximately independent samples simulated from the posterior distribution of the residual variance (σ), all singular values ($\lambda_1, \ldots, \lambda_{11}$) and the right and the left singular vector elements of genotypes and environments, respectively, whose 0.90 HPD and 0.95 HPD intervals do not contain the null value (0, 0). Data for grain yield measured in kilograms per hectare.

Parameter	Mean	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	0.90 HPD interval		0.95 HPD interval	
						Lower	Upper	Lower	Upper
σ	1311.00	39.13	1237.00	1310.00	1389.00	1247.40	1375.18	1234.90	1386.49
λ_1	2572.00	451.76	1683.00	2574.00	3446.00	1835.96	3323.21	1709.64	3468.89
λ_2	2354.00	388.40	1574.00	2364.00	3088.00	1730.65	3005.55	1585.22	3098.23
λ_3	2123.00	340.34	1429.00	2132.00	2763.00	1549.53	2662.10	1467.47	2794.07
λ_4	1815.00	328.85	1146.00	1823.00	2435.00	1277.49	2358.62	1167.92	2452.60
λ_5	1576.00	317.49	933.40	1586.00	2173.00	1058.67	2096.67	935.42	2173.17
λ_6	1262.00	311.66	630.20	1270.00	1847.00	758.68	1780.72	643.20	1856.85
λ_7	992.00	297.34	385.50	998.90	1555.00	506.17	1479.93	394.04	1562.37
λ_8	762.60	283.71	182.70	767.10	1308.00	299.76	1242.21	199.20	1321.56
λ_9	506.10	260.09	42.12	501.30	1027.00	15.90	865.98	0.04	949.55
λ_{10}	315.40	217.51	8.89	289.60	794.40	0.01	614.78	0.01	714.54
λ_{11}	167.40	153.72	1.82	124.30	553.80	0.00	388.81	0.00	472.82
$u_{12,2}$	0.3850	0.1966	−0.0371	0.4009	0.7242	0.0815	0.7176	0.0005	0.7524
$u_{8,3}$	−0.3714	0.1988	−0.7131	−0.3873	0.0618	−0.7054	−0.0679	−0.7464	0.0129
$u_{5,4}$	0.3671	0.2001	−0.0775	0.3838	0.7096	0.0490	0.6941	−0.0361	0.7358
$u_{1,6}$	0.3533	0.2163	−0.1170	0.3747	0.7167	0.0156	0.7073	−0.0702	0.7510
$v_{5,1}$	0.2878	0.1235	0.0389	0.2908	0.5203	0.0868	0.4911	0.0482	0.5271
$v_{9,1}$	−0.2769	0.1281	−0.5182	−0.2803	−0.0182	−0.4805	−0.0587	−0.5220	−0.0234
$v_{23,1}$	0.2779	0.1234	0.0276	0.2818	0.5106	0.0797	0.4846	0.0367	0.5167
$v_{25,1}$	0.4589	0.1104	0.2288	0.4637	0.6607	0.2804	0.6400	0.2431	0.6713
$v_{23,2}$	−0.3958	0.1268	−0.6263	−0.4017	−0.1312	−0.6090	−0.1960	−0.6338	−0.1419
$v_{7,3}$	0.3350	0.1323	0.0636	0.3402	0.5773	0.1221	0.5558	0.0758	0.5849
$v_{11,3}$	−0.2652	0.1319	−0.5161	−0.2668	0.0005	−0.4895	−0.0576	−0.5224	−0.0104
$v_{16,3}$	0.3249	0.1304	0.0547	0.3301	0.5674	0.1104	0.5372	0.0679	0.5755
$v_{24,4}$	−0.3918	0.1418	−0.6427	−0.4011	−0.0907	−0.6151	−0.1559	−0.6624	−0.1165
$v_{22,5}$	−0.3166	0.1581	−0.6007	−0.3248	0.0139	−0.5735	−0.0564	−0.6060	0.0055

Some description of the joint response of genotypes and environments can be given; for example, environments S5 and S25 and genotype 12 are located far from the center on the upper right-hand side of Figures 4(a) and 4(b), whereas genotype 2 is the farthest point from the center of the figure on the lower left-hand side, and is negatively related to genotype 12 and to environments S5 and S25. This result indicated that genotype 12 has significant positive interaction with S5, and S25, whereas genotype 2 has negative interaction with those environments. Concerning relationships among environments, S9 and S23 are located in opposite directions of the biplot [Figure 4(b)] and can be considered as two very different environments in

terms of genotype × environment interaction. Furthermore, environments S5, S23, and S25 do not overlap with environment S9; therefore they are also significantly different from environment S9 in terms of genotype × environment interaction variability. Since environments S5 and S25 do overlap with each other, they form a homogeneous group of environments but different from S9. Environments S5 and S25 did show some degree of similarity with S9.

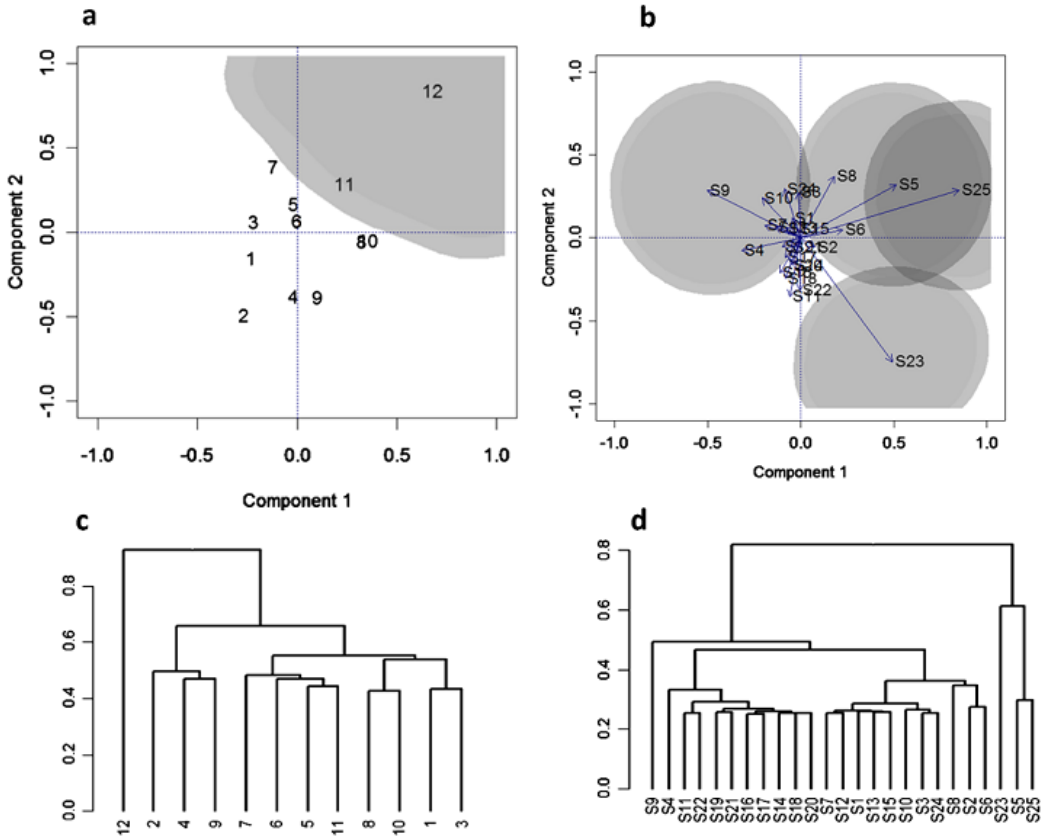


Figure 4. Plant breeding data of 12 genotypes in 25 environments for the first two components: (**a**) plot of bivariate the row scores $\mathbf{U}\mathbf{D}^{1/2}$ and the bivariate 0.95 (gray external contour) and 0.90 (gray internal contour) HPD regions; (**b**) plot of the bivariate column scores $\mathbf{V}\mathbf{D}^{1/2}$ and the bivariate 0.95 (gray external contour) and 0.90 (gray internal contour) HPD regions. Only genotype 12, and environments S5, S8, S9, S23, and S25 that do not include the null point (0, 0) at the 0.90 HPD probability level are depicted; (**c**) dendrogram of the 12 genotypes using the first two singular vectors; (**d**) dendrogram of the 25 environments using the first two right singular vectors.

Since the degree of overlap between significant environments based on the 0.90 HPD and 0.95 HPD regions differed and may not be clearly represented in the biplot, we performed a hierarchical cluster algorithm with a complete linkage strategy based on the posterior means of the Euclidean distance between the rows of the interaction matrices for genotypes and environments was performed; the dendrograms are presented in Figures 4(c) and (d), respectively. This complements what is depicted in Figures 4(a) and (b) very well. For example, genotype 12 forms a singleton opposite to the groups formed by the other genotypes [Figure 4(c)].

Concerning the environments, S5, S23, S25, and S9 are in opposite groups [Figure 4 (d)]. Furthermore, as already pointed out in the biplots of environments, environments S5, S23, and S25 formed a group in which each group member significantly contributes to genotype × environment variability; this group shows that S5 and S25 clustered earlier and S23 joined them later. Environments S5, S25 and S23 are different from environment S9. The overlapping of S5 and S25 is much more pronounced than the overlapping of S23; thus they joined early in the cluster. This approach for identifying subsets of homogeneous genotypes and environments that cause significant genotype × environment interaction is analogous (but not the same) to that presented by Burgueño et al. (2008) using a frequentist inference on a linear-bilinear model that belongs to the family of linear-bilinear models employed in this study.

Genotypes had longer HPD regions along the first and second bilinear terms than environments, indicating their large SD and great length of the HPD, which reflected the uncertainty of these estimates at both probability levels.

3.2.2 Implications for Breeding Trials of Bayesian Inference of Linear-Bilinear Models

The conditional posterior estimates of Bayesian linear-bilinear models for plant breeding data have the following advantages: (i) they provide a natural method for deriving confidence regions around the genotypic and environmental interaction parameters given by their scores, as represented in the biplot (and/or dendrogram); (ii) they facilitate the identification of genotypes and environments that cause significant

interaction and allow detecting groups of genotypes and environments with similar responses, (iii) they deal with unbalanced data (always present in plant breeding trials) in a natural manner; (iv) they can be used to efficiently incorporate information from historical plant breeding trials (prior) on environmental and genotypic means, or on dispersion parameters such as environmental, genotypic, or error variances; (v) they can be used naturally with unequal cell size; and (vi) they provide an efficient test for the significance of the number of GE bilinear component to be retained in the model.

Linear-bilinear models such as AMMI offer a family of models, rather than a single model; the general Bayesian computational methodology developed in this study can be applied to other linear-bilinear models by fixing certain parameters equal to 0 and relaxing some model constraints. For example, for $\boldsymbol{\alpha}=0$, $\mathbf{Y}=\mu\mathbf{1}_r\mathbf{1}_c'+\boldsymbol{\beta}'\otimes\mathbf{1}_r+\mathbf{UDV}'+\mathbf{E}$ is the column (site) regression model (SREG); for $\boldsymbol{\beta}=0$, $\mathbf{Y}=\mu\mathbf{1}_r\mathbf{1}_c'+\boldsymbol{\alpha}\otimes\mathbf{1}_c'+\mathbf{UDV}'+\mathbf{E}$ is the row (genotype) regression model (GREG); and for $\boldsymbol{\alpha}=\boldsymbol{\beta}=0$, $\mathbf{Y}=\mu\mathbf{1}_r\mathbf{1}_c'+\mathbf{UDV}'+\mathbf{E}$ is the complete multiplicative model (COMM). As the frequentist mixed-effect linear-bilinear model leads to a factor analytic structure of rows, columns and/or their interaction, the Bayesian paradigm of linear-bilinear models can be also represented in a factor analytic form.

This new approach offers new opportunities for efficiently incorporating historical data on environments and genotypes that should be useful for breeder's objectives, as well as forming density regions around the estimated interaction parameters. Furthermore, with the methodology presented in this article, meta-analysis (hierarchical analysis) involving year or other factors can be analyzed naturally. Although the computer time needed to process large plant breeding trials can be substantially greater than that needed to fit frequentist fixed or mixed linear-bilinear models, the continuous increase in computer power will minimize this disadvantage of the Bayesian estimation of linear-bilinear models over time. The Bayesian inference methodology described here is available in R in the following web page of CIMMYT.

http://www.cimmyt.org/english/wps/biometrics/index.htm.

4. CONCLUSIONS

In this research, we applied Bayesian inference for linear-bilinear models by using the multivariate von Mises-Fisher distribution as a prior for interaction parameters. In contrast to previous approaches, this estimation is not performed on the orthonormal eigenvectors but rather on the orthonormal matrices $\mathbf{U} = (\boldsymbol{u}_1, ..., \boldsymbol{u}_t)$ and $\mathbf{V} = (\boldsymbol{v}_1, ..., \boldsymbol{v}_t)$, and is done based on an MCMC sample from their posterior distribution; this satisfies model constraints and offers statistical inferential tools such as confidence regions for interaction parameters. Two data sets were used, one containing simulated data and one real plant breeding data. Results of the plant breeding trials show the usefulness of this general Bayesian approach for breeding trials and for detecting groups of genotypes and environments that cause interaction. For similar-structures data, this method is promising in that confidence regions for the GE interaction terms may be derived, unbalanced data are handled well, and the number of components to be retained in the model can be assessed by the Bayes factor. The Bayesian inference methodology described here can be extended to other linear-bilinear models by fixing certain parameters equal to zero and relaxing some model constraints.

APPENDIX A: THE VON MISES-FISHER DISTRIBUTION

The set of $r \times k$ orthonormal matrices is called the Stiefel manifold, which is denoted as $\mathcal{V}_{k,r}$. A probability distribution on $\mathcal{V}_{k,r}$, whose density has exponential form with linear and quadratic terms, is the matrix Bingham-von Mises-Fisher Distribution. The density function is given by

$$p(X|\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) \propto \text{etr}(\boldsymbol{C}'\boldsymbol{X} + \boldsymbol{B}\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X})$$

where **A** and **B** may be assumed symmetric and diagonal matrices, respectively. A random variable $\boldsymbol{X}$ with von Mises-Fisher distribution (Khatri and Mardia, 1977) is denoted as $\boldsymbol{X} \sim BMF(\boldsymbol{A} = \boldsymbol{0}, \boldsymbol{B} = \boldsymbol{0}, \boldsymbol{C})$. The normalization constant of the von Mises-Fisher density is given by the hypergeometric function of a matrix argument ${}_0\text{F}_1(\frac{1}{2}r, \frac{1}{4}D_\phi^2)$ where D_ϕ is the diagonal matrix of singular values of **C** (Herz, 1955; James, 1964) .

Table B.1. Posterior summary (Mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$) and 0.95 HPD intervals computed with 5000 approximately independent samples simulated from the joint posterior distribution of the 38 linear effects for plant breeding data of grain yield measured in kilograms per hectare.

Parameter	Mean	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	0.90 HPD interval		0.95 HPD interval	
						Lower	Upper	Lower	Upper
μ	3848.00	30.68	3788.00	3848.00	3909.00	3797.57	3898.59	3786.54	3907.27
α_1	135.90	102.92	−64.98	135.40	338.50	−25.06	312.32	−65.81	337.29
α_2	317.50	102.29	115.70	318.10	516.70	151.89	487.33	116.23	517.01
α_3	−64.68	103.26	−269.50	−64.43	134.80	−235.00	103.67	−265.60	138.61
α_4	2.75	102.62	−199.90	2.90	204.10	−157.74	178.77	−202.18	200.93
α_5	93.39	101.90	−106.50	93.61	292.90	−73.23	261.02	−106.95	292.18
α_6	181.00	101.81	−19.13	181.50	379.60	12.22	347.25	−16.59	381.85
α_7	499.30	103.21	297.00	499.20	704.60	329.03	668.15	291.42	697.90
α_8	64.30	102.72	−137.90	64.94	265.00	−99.60	237.96	−141.13	261.47
α_9	129.30	102.34	−72.10	129.60	328.40	−34.99	300.58	−75.84	324.12
α_{10}	−479.00	103.10	−682.50	−478.10	−278.20	−652.14	−311.59	−683.30	−279.28
α_{11}	−270.10	103.02	−471.50	−270.00	−66.31	−436.12	−99.78	−471.13	−66.11
α_{12}	−609.70	102.61	−809.80	−609.80	−405.00	−783.02	−447.03	−811.81	−407.89
β_1	134.50	151.21	−163.50	134.60	431.70	−109.56	386.25	−154.36	438.26
β_2	−533.60	151.39	−828.60	−532.60	−238.70	−784.48	−287.41	−831.18	−242.05
β_3	864.20	151.39	565.70	864.00	1163.00	611.24	1111.18	556.00	1150.01
β_4	−377.40	151.37	−677.60	−376.30	−76.97	−626.95	−130.58	−683.14	−84.26
β_5	1537.00	150.94	1240.00	1537.00	1837.00	1283.94	1782.23	1253.61	1848.75
β_6	−96.78	150.28	−390.40	−96.82	196.90	−344.40	150.31	−385.25	199.82
β_7	−1761.00	151.22	−2059.00	−1761.00	−1465.00	−2004.18	−1506.47	−2054.37	−1460.12
β_8	348.20	151.61	46.56	347.90	645.20	98.69	596.65	60.80	658.62
β_9	−1468.00	150.96	−1763.00	−1468.00	−1176.00	−1714.35	−1219.74	−1762.29	−1175.01

Table B.1. (*Continued.*)

Parameter	Mean	SD	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	0.90 HPD interval		0.95 HPD interval	
						Lower	Upper	Lower	Upper
β_{10}	−745.10	150.44	−1039.00	−745.00	−447.60	−997.52	−504.73	−1042.34	−452.14
β_{11}	−1202.00	151.64	−1497.00	−1203.00	−904.80	−1445.44	−949.06	−1496.57	−904.36
β_{12}	−1900.00	152.11	−2200.00	−1899.00	−1604.00	−2156.86	−1656.97	−2198.63	−1603.46
β_{13}	−971.60	150.75	−1268.00	−972.30	−676.10	−1217.06	−722.63	−1257.52	−668.61
β_{14}	−548.60	151.84	−851.10	−548.20	−252.70	−801.42	−301.01	−850.60	−252.51
β_{15}	−2215.00	151.62	−2513.00	−2214.00	−1920.00	−2461.53	−1962.21	−2517.25	−1925.26
β_{16}	87.58	150.95	−208.80	88.78	382.50	−158.01	338.84	−200.56	389.50
β_{17}	160.40	152.11	−138.50	161.60	460.70	−78.12	420.05	−138.58	460.65
β_{18}	1848.00	152.18	1546.00	1849.00	2148.00	1601.28	2102.67	1540.92	2140.93
β_{19}	−865.20	150.62	−1159.00	−865.60	−566.70	−1100.99	−607.88	−1161.24	−569.44
β_{20}	−1112.00	151.82	−1407.00	−1112.00	−810.00	−1362.40	−859.85	−1411.64	−815.98
β_{21}	−1627.00	151.76	−1930.00	−1627.00	−1332.00	−1878.18	−1378.41	−1929.72	−1331.88
β_{22}	4917.00	151.56	4621.00	4917.00	5212.00	4663.11	5158.02	4624.65	5214.83
β_{23}	2865.00	151.21	2570.00	2866.00	3160.00	2612.72	3110.06	2565.08	3155.47
β_{24}	911.80	149.76	614.40	911.10	1204.00	668.50	1158.86	625.06	1212.78
β_{25}	1749.00	150.27	1457.00	1747.00	2045.00	1492.73	1985.95	1456.93	2044.38

REFERENCES

Burgueño, J., J. Crossa, P. L. Cornelius, and R.-C. Yang. 2008. Using Factor Analytic Models for Joining Environments and Genotypes without Crossover Genotype × Environment Interaction. Crop Science doi:10.2135/cropsci2007.11.0632.

Chib, S. "Marginal Likelihood From the Gibbs Output," Journal of the American Statistical Association, 90, 1313-1321.

Cornelius, P.L., M.S. Seyedsadr, and J. Crossa. 1992. Using the shifted multiplicative model to search for 'separability in crop cultivar trials. Theoretical and Applied Genetic. 84:161-172.

Cornelius, P.L., J. Crossa, and M.S. Seyedsadr. 1994. Tests and estimators of multiplicative models for variety trials. Proceedings of the 1993 Kansas State University Conference on Applied Statisticas in Agriculture. Manhattan, Kansas.

Cornelius, P.L., J. Crossa, and M.S. Seyedsadr. 1996. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. p. 199-234. In M.S. Kang and H.G. Gauch (eds.). Genotype-by-environment interaction. CRC Press, Boca Raton, FL.

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

Crossa, J., R.-C. Yang, 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.

Crossa, J., S. Perez-Elizalde, D. Jarquin, J. M. Cotes, K. Viele, G. Liu, and P.L. Cornelius. 2011. Bayesian estimation of the Additive Main effects and Multiplicative Interaction (AMMI) model. Crop Science, 51, 1458-1469.

Gabriel, K.R. 1978. Least squares approximation of matrices by additive and multiplicative models. Journal of the Royal Statistical Society, Series B, 40:186-196.

Gauch, H.G. 1988. Model selection and validation for yield trials with interaction. Biometrics 44:705-715.

Gelman, A., and D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457-511.

Gollob, H.F. 1968. A statistical model which combines features of factor analytic and analysis of variance. Psychometrika 33:73-115.

Han, C. and Carlin, B. P. (2001). Markov Chain Monte Carlo Methods for Computing Bayes Factors: A comparative Review. Journal of the American Statistical Association, 96: 1122-1132.

Herz, C. S. (1955). Bessel functions of matrix argument. Annals of Mathematics, 61:474–523.

Hoff, P.D. 2009. Simulation of the matrix Bingham-von Mises-Fisher distribution, with applications to multivariate and relational data. Journal of Computational and Graphical Statistics 18:438-456.

James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. Annals of Mathematical Statistics, 35: 475-501

Jeffreys, H. (1961). Theorey of Probability, Oxford: Clarendon Press.

Johnson, D.E., and G.A. Graybill. 1972. An analysis of a two-way model with interaction and no replication. Journal of the American Statistical Association 67:862-868.

Kass, R.E. and Raftery, A. (1995). Bayesian factors. Journal of the American Statistical

Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. J. Agric. Sci. 103:123-135.

Khatri, C.G., and K.V. Mardia. 1977. The von Mises-Fisher matrix distribution in orientation statistics. Journal of the Royal Statistical Society. Series B (Methodological) Vol. 39 No. 1:95-106.

Koev, P. and Edelman, A. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. Mathematics of Computation, 75: 833-845.

Liu, G. 2001. Bayesian computation for general linear-bilinear models. Ph.D. dissertation. Department of Statistics. University of Kentucky. Lexington, KY.

Mandel, J. 1969. The partitioning of interaction in analysis of variance. Journal of Research of the National Bureau of Standards, Series B, 73:309-328.

Mandel, J. 1971. A new analysis of variance models for non-additive data. Technometrics 13:1-18.

Marasinghe, M.G. 1985. Asymptotic tests and Montecarlo studies associated with the multiplicative interaction model. Communication in Statistics A. Theory and Methods 14:2219-2231.

Mardia, K.V., J.T. Kent, and J.B. Bibbi. 1979. Multivariate analysis. Academic Press: London, UK.

Raftery, A.E., and S.M. Lewis. 1995. The number of iterations, convergence diagnostics and generic Metropolis algorithms. *In* Practical Markov Chain Monte Carlo. W.R. Gilks, D.J. Spiegelhalter, and S. Richardson (eds.). London, U.K.: Chapman and Hall.

Schott, J.R. 1986. A note on the critical value in stepwise tests of multiplicative components of interactions. Communication in Statistics A. Theory and Methods 15:1561-1570.

Seyedsadr, M.S., and P.L. Cornelius. 1992. Shifted multiplicative models for non-additive two-way tables. Communications in Statistics, B Simulation and Computation 21:807-822.

Viele, K., and C. Srinivasan. 2000. Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback-Leibler information. Journal of Statistical Planning and Inference 84:201-219.

CAPÍTULO 3: META-ANALYSIS OF PLANT BREEDING TRIALS USING BAYESIAN HIERARCHICAL MODELS

Diego Jarquín, Sergio Perez-Elizalde, José Crossa, and Ky Mathews

RESUMEN

En el fitomejoramiento, se establecen experimentos en múltiples ambientes y múltiples años para evaluar y predecir el desempeño de los genotipo bajo diferentes condiciones ambientales; y para cuantificar, estudiar e interpretar la interacción genotipo x ambiente (GE) de forma que los mejores genotipos sean seleccionados, recombinados y plantados en los años subsecuentes y en otros ambientes. Este tipo de datos se registran en tablas de doble entrada y los parámetros de interés para la inferencia estadística son los efectos bilineales de GE o una combinación de GE mas algún efecto lineal (genotipos o ambientes). Desde una perspectiva bayesiana la inferencia sobre los parámetros bilineales de la interacción se basa en la distribución posterior de las matrices ortonormales U y V que resultan de la descomposición en valores singulares (SVD) $\mathbf{UDV}'$ de la matriz de parámetros interacción GE. En este documento se propone un modelo jerárquico bayesiano para experimentos de fitomejoramiento provenientes de varios ambientes y años. Para U y V se considera la distribución matricial von-Mises Fisher (mVMF); para tales parámetros una a priori matemáticamente conveniente es una distribución conjugada condicional mVMF. Para los efectos lineales se utiliza la estructura normal jerárquica y los parámetros de precisión se considera siguen una distribución gamma. Debido a la alta dimensión del espacio paramétrico la computación de la distribución posterior conjunta de los parámetros del modelo se realiza a través de MCMC. Un conjunto de datos de trigo de un experimento de múltiples ambientes para tres años consecutivos fue usado como ilustración, el conjunto de datos del primer año fue usado como información a priori. Los resultados muestran que el modelo propuesto permite la identificación de grupos de genotipos y ambientes que causan la interacción GE.

Palabras clave: Interacción de doble entrada, modelo de efectos principales aditivos e interacción multiplicativa (AMMI).

Capítulo en revisión para su publicación.

ABSTRACT

In plant breeding, multienvironment trials in multiple years are established to evaluate and predict genotype performance under different environmental conditions, and to quantify, study, and interpret genotype × environment interaction (GE) such that the best genotypes are selected, recombined, and planted in subsequent years and other environments. This kind of data is accommodated in a two-way table and the parameters of inferential interest are the bilinear GE effects or a combination of GE plus some of the linear effects (genotypes or environments). From the Bayesian perspective, the inference over the bilinear interaction parameters is based on the posterior distribution of the orthonormal matrices **U** and **V** that result from the singular value decomposition (SVD) $\mathbf{UDV}'$ of the GE interaction matrix of parameters. In this paper we propose a Bayesian hierarchical model for plant breeding trials data arising from several environments and years. For **U** and **V** the matrix von-Mises Fisher (mVMF) distribution is considered; for its parameters a mathematically convenient priors are conditional conjugate mVMF distributions. For the linear effects the usual normal hierarchical structure is utilized and priors for precisions are assumed to follow gamma distributions. Due to the high dimensional parametric space the computation of the joint posterior distribution of the model parameters is made through MCMC. A wheat multi environment data set for three consecutive trials was used as illustration; the data set from the first year was used as prior information. This example depicts the main features of our proposed model for the analysis of GE interaction effects, namely the isolation of interaction effect from the inter trial variation. Also, the results show that the proposed model allows the identification of groups of genotypes and environments that cause GE interaction.

Key words: Two way interaction, Additive Main Effect and Multiplicative Interaction (AMMI) model

INTRODUCTION

Models combining linear (additive) and bilinear (multiplicative or non-additive) terms are useful for the analysis of two-way tables with interaction (Cornelius and Seyedsadr, 1997). This is particularly important in agriculture and plant breeding where genotypes are planted in environments (locations and years) and the genotype × environment interaction (GE) needs to be studied for selection decisions to be carried out in order to assemble the genotypes for the next cycle of a breeding program.

The usual two-way analysis of variance model for *r* rows and *c* columns is

$$\bar{y}_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \bar{\varepsilon}_{ij} \qquad (1)$$

where μ, α_i, β_j, and $(\alpha\beta)_{ij}$ (for i=1,2,…,r; and j=1,2,…,c,) are the grand mean, the effect of the ith row (genotypes), the effect of the jth column (environments) , and the effect of the interaction of the ith row on the jth column, respectively. The $\bar{\varepsilon}_{ijl}$ are identically and independently distributed with $N(0,\sigma_e^2/n_{ij})$ (for simplicity in what follows we supposed equal number of observations n in each cell). The linear terms (μ, α_i, and β_j) of the linear-bilinear models are first fitted by ordinary least square and the bilinear term $(\alpha\beta)_{ij}$ is fitted through the singular value decomposition (SVD) after fitting linear effects (Gabriel, 1978). This yields the usual linear-bilinear two-way model (Gollob,1968; Mandel, 1969, 1971) that is used in plant breeding trials for assessing adaptation and stability (Kempton, 1984; Gauch, 1988; Cornelius et al., 1996; Crossa et al., 2004) and it is named the Additive Main effect and Multiplicative Interaction (AMMI) model (Gauch, 1988)

$$\bar{y}_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk} + \bar{\varepsilon}_{ij} \qquad (2)$$

where, λ_k is the singular value of SVD of subject to $\lambda_1 \geq ... \lambda_t \geq 0$; u_{ik} and v_{jk} are the left and right singular vectors, respectively with the constraints that $\sum_i u_{ik}^2 = \sum_j v_{jk}^2 = 1$ and, for $k \neq k'$, $\sum_i \alpha_{ik}\alpha_{ik'} = \sum_j \beta_{jk}\beta_{jk'} = 0$ and $t = min(r,c) - 1$.

In matrix notation (2) can be expressed as

$$\mathbf{Y} = \mu \mathbf{1}_r \mathbf{1}_c' + \boldsymbol{\alpha} \otimes \mathbf{1}_c' + \boldsymbol{\beta}' \otimes \mathbf{1}_r + \mathbf{UDV}' + \mathbf{E} \qquad (3)$$

where $\mathbf{Y} = [\bar{y}_{ij}]$, $\boldsymbol{\alpha} = [\alpha_i]$, $\boldsymbol{\beta} = [\beta_j]$, $\mathbf{D} = \text{diag}(\lambda_k, k = 1,2,...t)$, $\mathbf{U} = (\boldsymbol{u}_1,...,\boldsymbol{u}_t)$, $\boldsymbol{u}_k = [u_{ik}]$ $\mathbf{V} = (\boldsymbol{v}_1,...,\boldsymbol{v}_t)$, $\boldsymbol{v}_k = [v_{jk}]$, and $\mathbf{E} = [\bar{\varepsilon}_{ij}]$. Linear-bilinear models such as that in (3) offers a family of models, for example, for $\boldsymbol{\alpha} = \mathbf{0}$, $\mathbf{Y} = \mu\mathbf{1}_r\mathbf{1}_c' + \boldsymbol{\beta}' \otimes \mathbf{1}_r + \mathbf{UDV}' + \mathbf{E}$ is the column regression model; for $\boldsymbol{\beta} = \mathbf{0}$, $\mathbf{Y} = \mu\mathbf{1}_r\mathbf{1}_c' + \boldsymbol{\alpha} \otimes \mathbf{1}_c' + \mathbf{UDV}' + \mathbf{E}$ is the row regression model; and for $\boldsymbol{\alpha} = \boldsymbol{\beta} = \mathbf{0}$, $\mathbf{Y} = \mu\mathbf{1}_r\mathbf{1}_c' + \mathbf{UDV}' + \mathbf{E}$ is the complete bilinear (multiplicative) model (COMM). In plant breeding terminology the genotypes denote the rows and the environments (site) the columns; therefore the column regression model is named the Site Regression (SREG) (Crossa and Cornelius, 1997), and the row regression model is referred to the Genotype Regression (GREG). When allowing $\boldsymbol{\alpha} = \mathbf{0}$ (and/or $\boldsymbol{\beta} = \mathbf{0}$) this linear effect is absorbed in the bilinear decomposition $\mathbf{UDV}'$ such that the main effects of row (or column) is estimated in combination with the GE effect.

With the least square method the parameters in (3) are estimated by fitting first the linear terms ignoring the bilinear terms that are subsequently fitted as the first *t* components of the singular value decomposition of the residual matrix $\mathbf{Z} = \mathbf{Y} - \hat{\mu}\mathbf{1}_r\mathbf{1}_c' - \hat{\boldsymbol{\alpha}} \otimes \mathbf{1}_c' - \hat{\boldsymbol{\beta}}' \otimes \mathbf{1}_r$ where $\hat{\mu}$, $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are the LS estimates obtained in the first step, (Gabriel, 1978).

Viele and Srinivasan (2000) proposed Bayesian estimation of parameters for model (3) using a spherical uniform prior distributions for the bilinear effects and the posterior means as shrinkage estimates. The unpublished PhD thesis of G. Liu (Liu, 2001), used the same prior distributions as Viele and Srnivasan (2000), and derived the posterior full conditional distributions of unknowns in model (3) such that a Gibbs sampling of the joint posterior distribution could be used. The approach of Viele and Srinivasan (2000) for sampling the conditional posterior distributions was performed within the vector framework; that is, for the joint posterior distribution of the columns of $\boldsymbol{u}_k$ and $\boldsymbol{v}_k$ which are the columns of **U** and **V**., respectively. Recently, Crossa et al. (2011) used this approach to analyze real data of plant breeding multi-environment trials; the authors showed that inferential statistics can be naturally incorporated by adopting the Bayesian approach, as well as estimation of the GE interaction parameters, joint credible regions for phenotypic and genotypic scores are easily obtained by using a MCMC sample from the joint posterior.

A generalization of the vector approach to a matrix approach for the Bayesian inference of model (3) was proposed by Perez-Elizalde et al. (2011) who used von Mises-Fisher distributions as priors for the orthonormal matrices **U** and **V**. To sample from the von Mises-Fisher distribution onto the multi-dimensional sphere, the authors used the algorithm proposed by Hoff (2007) to generate samples from the posterior distributions of orthonormal matrices that arise in the analysis of multivariate data.

Meta-analysis is a useful tool for summarizing and integrating the findings of several research studies. As a method for combining information from several parallel data sources, meta-analysis is closely connected to Bayesian hierarchical modeling (Gelman 2004). The Bayesian approach gives further advantages over other approaches because it is possible to formally take into account data sets from previous experiments or from the researchers' expertise. Plant breeders often perform analysis of two-way tables with the aim of finding groups of genotypes and environments with GE interaction effect on a phenotypic trait. Usually breeders have historical records of the experimental data that can be incorporated in a meta-analysis involving several trials in different environments and years.

In this chapter we extend the method of Perez-Elizalde et al. (2011) of the Bayesian analysis of model (3) (Bayesian AMMI) with the objective of developing this analysis in a hierarchical modeling framework. We explain the hierarchical Bayesian inference of model (3) to real plant breeding multi environment trial comprising 12 genotypes and 25 environments evaluated in two consecutive years; this data was also used by Perez-Elizalde et al., 2011 that fitted the Bayesian AMMI model. As conditional conjugate priors for the orthonormal matrices produced by the singular value decomposition of the interaction matrices, the von Mises-Fisher distribution is used. Bivariate highest probability density regions (HPD) were estimated for the posterior distributions of the first two phenotypic and genotypic scores which are scaled bilinear interaction parameters.

THE HIERARCHICAL BAYESIAN AMMI MODEL

The model proposed by Perez-Elizalde et al (2011) is an extension of the model proposed by Viele and Srinivasan (2000) and Crossa et al (2011). In Perez-Elizalde's Bayesian inference of model (3) a multivariate normal prior is assumed for the linear effects $\boldsymbol{\theta} = (\mu, \boldsymbol{\alpha}, \boldsymbol{\beta})$, and the priors for the orthonormal matrices $(\mathbf{U}, \mathbf{V})$ are matrix von Mises-Fisher (mVMF) distributions while the elements of the diagonal matrix **D** follow a truncated normal distribution. Hoff (2009) developed a method for sampling from the von Mises-Fisher distribution on the multi-dimensional sphere. In this paper we use the model above to construct a hierarchical model.

The hierarchical model we proposed has two levels, the first level corresponds to experimental data in each evaluation time (year) across the corresponding environments (i.e., each year corresponds to a population) while the second level is indexed by the parameters of the populations from where the observations came, i.e., this model considers each experiment as a realization of a super population model.

THE FIRST LEVEL MODEL

Suppose we have h periods (year)of evaluation , then for the m^{th} data set, $m = 1, \ldots, h$, and following (3) the model is

$$\mathbf{Y}_m = \mu_m \mathbf{1}_r \mathbf{1}_c^{'} + \boldsymbol{\alpha}_m \otimes \mathbf{1}_c^{'} + \boldsymbol{\beta}_m^{'} \otimes \mathbf{1}_r + \mathbf{U}_m \mathbf{D}_m \mathbf{V}_m^{'} + \mathbf{E}_m \qquad (4)$$

In this model the parameters remains as in the model (3) and are identified for each evaluation time (year) by the m suffix. Additionally, suppose that the parameters $\{\mu_m, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{V}_m\}_{m=1}^{h}$ which index the distributions of $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_h$, $m = 1, \ldots, h$, form an exchangeable sequence whose underlying distribution is given below.

SECOND LEVEL MODEL

For the linear terms $\boldsymbol{\theta}_m = (\mu_m, \boldsymbol{\alpha}_m, \boldsymbol{\beta}_m)$ in model (4), we assume a (1 + *r* + *c*) multivariate normal distribution with mean $\boldsymbol{\theta} = (\mu, \boldsymbol{\alpha}, \boldsymbol{\beta})$ and singular covariance matrix

$$\boldsymbol{\Sigma}_m = n_m \tau_m \begin{bmatrix} (r * c)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (c)^{-1}\mathbf{K}_r\mathbf{K}_r^{'} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (r)^{-1}\mathbf{K}_c\mathbf{K}_c^{'} \end{bmatrix}$$

where $\mathbf{K}_s$ is a matrix such that $\mathbf{K}_w'\mathbf{K}_w = \mathbf{I}_{w-1}$ and $\mathbf{K}_w\mathbf{K}_w' = \mathbf{I}_w - \frac{1}{w}\mathbf{J}_w$ and $\mathbf{J}_w$ is an $s \times s$ matrix with all elements equal to one. Due to the restrictions $\boldsymbol{\alpha}'\mathbf{1}_r = 0$ and $\boldsymbol{\beta}'\mathbf{1}_c = 0$ it is required a one to one transformation such as $(\boldsymbol{\alpha}_m^*, \boldsymbol{\beta}_m^*) = (\mathbf{K}_r'\boldsymbol{\alpha}_m, \mathbf{K}_c'\boldsymbol{\beta}_m)$.

Let $\boldsymbol{\theta}_m^* = (\mu_m, \boldsymbol{\alpha}_m^*, \boldsymbol{\beta}_m^*)$, then its prior density function is

$$\pi\left(\boldsymbol{\theta}_m^* \mid \Sigma_m^*\right) \propto \left|\Sigma_m^*\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}_m^* - \boldsymbol{\theta}_m\right)' \Sigma_m^{*-1}\left(\boldsymbol{\theta}_m^* - \boldsymbol{\theta}_m\right)\right\} \qquad (5)$$

which corresponds to a multivariate normal distribution with mean $\boldsymbol{\theta}^*$ and block diagonal covariance matrix given by

$$\boldsymbol{\Sigma}_m^* = n_0 \tau_m \begin{bmatrix} (r_0 * c_0)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (c_0)^{-1}\mathbf{I}_{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (r_0)^{-1}\mathbf{I}_{c-1} \end{bmatrix}$$

where $\boldsymbol{\theta}^* = (\mu\ , \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ correspond to the prior vector of means of the main effects.

As shown in Pérez-Elizalde et al (2011), given $\boldsymbol{\theta}_m$ and τ_m, the conditional likelihood function for the matrices $(\mathbf{U}_m, \mathbf{D}_m, \mathbf{V}_m)$ is

$$\begin{aligned} L(\mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m \mid \boldsymbol{\theta}_m, \tau_m, \mathbf{Y}_m) = L(\mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m \mid \tau_m, \mathbf{Y}_m) &\propto \exp\{-\tfrac{n_m\tau_m}{2}\operatorname{tr}((-2\mathbf{Y}_m + \mathbf{U}_m\mathbf{D}_m\mathbf{V}_m^{'})(\mathbf{U}_m\mathbf{D}_m\mathbf{V}_m^{'})^{'})\} \\ &= \operatorname{etr}\{-\tfrac{n_m\tau_m}{2}(-2\mathbf{Y}_m + \mathbf{U}_m\mathbf{D}_m\mathbf{V}_m^{'})(\mathbf{U}_m\mathbf{D}_m\mathbf{V}_m^{'})^{'}\} \qquad (6) \end{aligned}$$

where "etr" denote the exponential of the trace. Then, mathematically convenient conditional distributions for $\mathbf{U}_m$ and $\mathbf{V}_m$ are of the form

$$\pi(\mathbf{U}_m \mid \mathbf{V}, \mathbf{D}, \tau_m, \mathbf{M}_{0m}) \propto \operatorname{etr}(n_0\tau_m\mathbf{M}_{0m}\mathbf{V}\mathbf{D}\mathbf{U}_m') \qquad (7)$$

and

$$\pi(\mathbf{V}_m \mid \mathbf{U}, \mathbf{D}, \tau_m, \mathbf{M}_{0m}) \propto \text{etr}(n_0 \tau_m \mathbf{M}'_{0m} \mathbf{U}\mathbf{D}\mathbf{V}'_m) \tag{8}$$

which are mVMF distributions.

For easy handling we consider for each one of the diagonal elements of $\mathbf{D}_m$, $\lambda_{1_m} > \lambda_{2_m} > ... > \lambda_{t_m}$, a conditional conjugate left truncated normal distributions with density given by

$$\pi(\lambda_{k_m} \mid \tau_m, l_k) = \left\{1 - \Phi\left(\sqrt{n_0 \tau_m}(\lambda_{(k+1)_m} - l_k)\right)\right\}^{-1} \mathrm{N}\left(\lambda_{k_m} \mid l_k, (n_0 \tau_m)^{-1}\right), \tag{9}$$

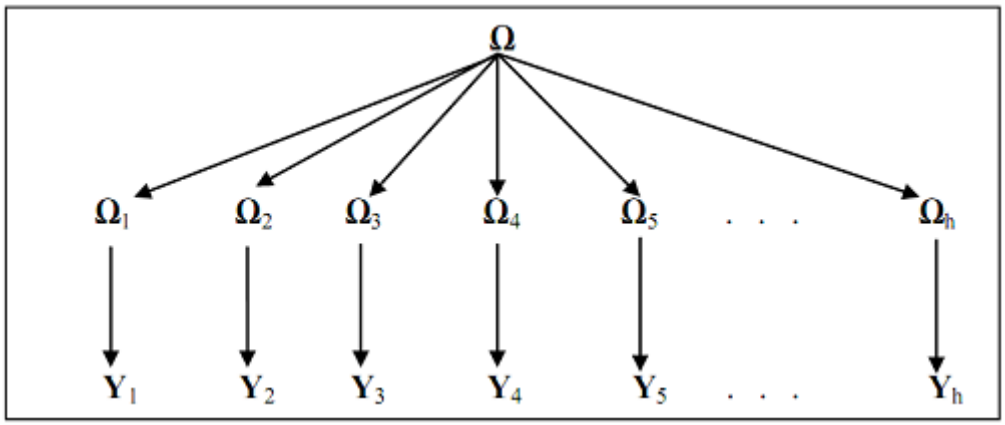
$k = 1, ..., t, \lambda_{(t+1)_m} = 0$.

where $l_1, l_2, ..., l_t$ correspond to the diagonal elements of **D**.

For the precision parameter τ_m the joint likelihood in (11) suggest a conjugate prior following a gamma distribution with parameters $a/2$, and $S_{0m}^2/2$; that is

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

The general structure of the hierarchical model may be represented graphically as follows



where for each period of evaluation we have $\mathbf{\Omega}_m = \{\boldsymbol{\theta}_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{V}_m, \tau_m\}$; and the super-population parameters are denoted by $\mathbf{\Omega} = \{\boldsymbol{\theta}, \mathbf{U}, \mathbf{V}, \mathbf{D}\}$.

LIKELIHOOD FUNCTION

The likelihood function for parameters of model (3) is given by

$$\mathbf{L}\left\{\left(\boldsymbol{\theta}_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \tau_m \mid \mathbf{Y}_m\right)\right\}_{m=1}^{h} \propto \prod_{m=1}^{h} \tau^{\frac{n_m r_m c_m}{2}} \exp\{\frac{\tau_m}{2}[n_m \operatorname{tr}(\mathbf{E}_m \mathbf{E}_m^{'}) - (n_m - 1)\operatorname{tr}(\mathbf{S}_m \mathbf{S}_m^{'})]\} \qquad (11)$$

where

$\tau_m = 1/\sigma_m^2,$

$\mathbf{S}_m = \left\{\sqrt{s_{mij}^2}\right\},$

$$s_{ijm}^2 = \frac{\sum_{l=1}^{n} (\bar{\mathbf{Y}}_{ijm} - Y_{ijlm})^2}{n_m - 1}$$

$$\mathbf{E}_m = \mathbf{Y}_m - \mu_m \mathbf{1}_r \mathbf{1}_c^{'} - \boldsymbol{\alpha}_m \otimes \mathbf{1}_c^{'} - \boldsymbol{\beta}_m^{'} \otimes \mathbf{1}_r - \mathbf{U}_m \mathbf{D}_m \mathbf{V}_m^{'}.$$

We used the functional form from (5-10) to define conditional conjugate priors for the first level model parameters. As will be shown (12-15), with this criterion we obtain priors distributions which are easily elicit able with information available to most plant breeders.

THE PRIOR DISTRIBUTION

In order to work with a hierarchical model we assign prior distributions to the super-population parameters $(\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{D}, \mathbf{U}, \mathbf{V})$, for the linear terms $\boldsymbol{\theta}^* = (\mu, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*)$ a conjugate prior density is

$$\pi\left(\boldsymbol{\theta}^* \mid \Sigma_p^*\right) \propto \left|\Sigma_p^*\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0^*\right)' \Sigma_p^{*-1}\left(\boldsymbol{\theta}^* - \boldsymbol{\theta}_0^*\right)\right\} \qquad (12)$$

$$\boldsymbol{\Sigma}_p^* = \tau_0 \begin{bmatrix} 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{c-1} \end{bmatrix}$$

where $\boldsymbol{\theta}_0^* = (\mu_0, \mathbf{K}_r' \boldsymbol{\alpha}_0, \mathbf{K}_c' \boldsymbol{\beta}_0)$.

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

$$\pi(\mathbf{U} \mid \mathbf{V}_0, \mathbf{D}_0, \tau_0, \mathbf{M}_0) \propto \operatorname{etr}(\tau_0 \mathbf{M}_0 \mathbf{V}_0 \mathbf{D}_0 \mathbf{U}') \qquad (13)$$

and

$$\pi(\mathbf{V} \mid \mathbf{U}_0, \mathbf{D}_0, \tau_0, \mathbf{M}_0) \propto \text{etr}(\tau_0 \mathbf{M}_0' \mathbf{U}_0 \mathbf{D}_0 \mathbf{V}') \tag{14}$$

In the above priors we can introduce available prior information thorough their hyperparameters. For example, $\mathbf{U}_0$, $\mathbf{D}_0$ and $\mathbf{V}_0$ could be seen as the singular value decomposition of

$$\mathbf{Z}_0 = \mathbf{M}_0 - \mu_0 \mathbf{1}_r \mathbf{1}_c^{'} - \boldsymbol{\alpha}_0 \otimes \mathbf{1}_c^{'} - \boldsymbol{\beta}_0^{'} \otimes \mathbf{1}_r$$

such that $(l_{01}, l_{02}, ..., l_{0t}) = diag(\mathbf{U}_0^{'} \mathbf{Y}_0 \mathbf{V}_0)$, where $\mathbf{M}_0$ is the matrix of prior predicted cell means; μ_0, $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ are prior beliefs about the linear row and column effects.

For the elements of the diagonal matrix **D** of singular values we have as a conjugate priors left truncated normal distributions with density

$$\pi(\lambda_k \mid \tau_0) = \left\{1 - \Phi\left(\sqrt{\tau_0}(\lambda_{k+1} - l_k^0)\right)\right\}^{-1} \text{N}\left(\lambda_k \mid l_k^0, (\tau_0)^{-1}\right), \quad k = 1, ..., t, \quad \lambda_{t+1} = 0 \tag{15}$$

In the equation above the hyperparameter l_k^0 is the prior mean and may be elicitated thorough the singular value decomposition already described.

THE POSTERIOR DISTRIBUTION

The joint posterior distribution is obtained following the Bayes' rule as the product of the period of evaluation likelihoods in (11), the joint second level distribution given by (5, 7-10) and the joint prior obtained as the product of densities (12-15). That is, the joint posterior is given by

$$\prod_{m=1}^{h} L(\boldsymbol{\theta}_m^*, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \tau_m \mid \mathbf{Y}_m) \pi(\boldsymbol{\theta}_m^*, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \tau_m \mid \boldsymbol{\theta}^*, \mathbf{U}, \mathbf{D}, \mathbf{V}) \pi(\boldsymbol{\theta}^*, \mathbf{U}, \mathbf{D}, \mathbf{V}) \tag{16}$$

Evidently, the joint posterior (16) is high dimensional so a MCMC technique is necessary in order to obtain estimations of marginal posterior distributions and their summaries. In this case, due to conditional conjugate structure of the involved distributions, the Gibbs sampling is an appropriate strategy to sample from the posterior. As it is well known, the first step of the Gibbs

sampler algorithm is to determine the full conditionals. From (16), it is straightforward to derive the posterior conditionals which are in Appendix A.

Sampling sequentially from these conditionals, after a large number of iterations, we obtain a sample from the joint posterior distribution which it may use to calculate posterior summaries i.e., means, standard deviations, and HPD intervals.

THE GIBBS SAMPLER

The Gibbs Sampling is implemented in two stages; the first stage consists of drawing samples from the joint-super population parameters through their full conditional posterior distributions, in the second stage the samples are obtained from the full conditional posterior distributions of the parameters in each trial data set.

The Gibbs sampling scheme proceed in the following order. For the m^{th} period (m=1,…, h) and for a sample of size s, simulate from:

$$\boldsymbol{\theta}^{*(i+1)} \sim N_{r+c-1}\left(\boldsymbol{\theta}^{*}\middle|\ (\sum_{m=1}^{h}\sum_{m}^{*\,-1}+\sum_{p}^{*\,-1})^{-1}\ \left(\sum_{m=1}^{h}\sum_{m}^{*\,-1}\widehat{\boldsymbol{\theta}}_{m}^{*}+\sum_{p}^{*\,-1}\boldsymbol{\theta}_{0}^{*}\right),(\sum_{m=1}^{h}\sum_{m}^{*\,-1}+\sum_{p}^{*\,-1})^{-1}\right)$$

$$\mathbf{D}^{(i+1)} \sim \pi(\mathbf{D}\,|\,\tau_m,\mathbf{D}_m)$$

$$\mathbf{U}^{(i+1)} \sim \pi(\mathbf{U}\,|\,\tau_m,\mathbf{V}_m,\mathbf{M}_{0m},\mathbf{D})$$

$$\mathbf{V}^{(i+1)} \sim \pi(\mathbf{V}\,|\,\tau_m,\mathbf{U}_m,\mathbf{M}_{0m},\mathbf{D})$$

After a sample from the conditional posteriors of the first stage parameters has been generated, a sample from the conditional of the second stage parameters must be simulated in the order indicated below

$$\tau_m^{(i+1)} \sim \pi(\tau_m\,|\,a_{\tau_m},b_{\tau_m},\mathbf{E}_m)$$

$$\boldsymbol{\theta}_m^{*(i+1)} \sim N_{r+c-1}\left(\boldsymbol{\theta}_m^{*}\middle|\ (\sum_{a}^{*\,-1}+\sum_{m}^{*\,-1})^{-1}\ \left(\sum_{a}^{*\,-1}\widehat{\boldsymbol{\theta}}_{m}^{*}+\sum_{m}^{*\,-1}\boldsymbol{\theta}^{*}\right),(\sum_{a}^{*\,-1}+\sum_{m}^{*\,-1})^{-1}\right)$$

$$\mathbf{D}_m^{(i+1)} \sim \pi(\mathbf{D}_m\,|\,\tau_m,\mathbf{U}_m,\mathbf{V}_m,\mathbf{Y}_m,\mathbf{D})$$

$$\mathbf{U}_m^{(i+1)} \sim \pi(\mathbf{U}_m\,|\,\tau_m,\mathbf{V}_m,\mathbf{D}_m,\mathbf{Y}_m,\mathbf{M}_{0m},\mathbf{V},\mathbf{D})$$

$$\mathbf{V}_m^{(i+1)} \sim \pi(\mathbf{V}_m\,|\,\tau_m,\mathbf{U}_m,\mathbf{D}_m,\mathbf{Y}_m,\mathbf{M}_{0m},\mathbf{U},\mathbf{D})$$

RESULTS AND DISCUSSION

PLANT BREEDING DATA

The multi-environment multi-year plant breeding trial data analyzed in this section comprises 20 wheat varieties evaluated in 12 environments for three consecutive years. The first-year data were used to elicit the prior for the super population model. The posterior means of the 33 linear parameters (overall mean, 20 genotypic effects and 12 environmental effects) and the 0.95 HPD intervals corresponding to super population model and each period of evaluation (years) model are given in Table B1, Table B2 and Table B3 (Appendix B), respectively.

Given in Table 1, Table 2 and Table 3 for the super population parameters and for the individual year parameters are the means, standard deviation quartiles and 0.95 HPD intervals for λ_1- λ_2, some elements of the singular vectors of genotypes (u_{ik}) and the singular vector of environments (v_{jk}) whose 0.95 HPD do not contain the null value.

In general, for both cases (the super population model and the individual year model) the absolute values of u_{i1} and v_{j1} are larger than those values of u_{i2} and v_{j2}, whereas the standard deviation values (SD) for u_{i1} and v_{j1} are smaller than the corresponding values for u_{i2} and v_{j2}. Thus, there is more uncertainty in the second bilinear component for genotypes and environments than in the first component; these results are in agreement with those obtained by Perez-Elizalde et al (2011).

CONFIDENCE REGIONS OF THE FIRST TWO BILINEAR TERMS

Hierarchical analysis

The HPD regions of the genotype and environmental scores that have high probability of being different from the null point (0, 0) can be seen in Table 1 and in the biplot depicted in Figure 1a and Figure 3a. Figure 1a shows the plot of the genotype scores $\mathbf{U'D}^{1/2}$, where the outer and inner shaded areas of the graph are the bivariate 0.95 and 0.90 HPD posterior regions respectively, for the scores of genotypes 1, 2, and 5. The scores $u_{1,1}, u_{2,1}, u_{5,1}, v_{6,1}$ and $v_{11,1}$ are those that do not include the null point (0, 0) for the bivariate 0.95 HPD region and thus

cause most of the significant interaction. Other posterior 0.95 HPD intervals for eigenvector $u_{i,1}$ and $u_{i,2}$ or $v_{j,1}$ and $v_{j,2}$ elements might not contain zero, but the 0.95 bivariate HPD region for their corresponding scores $(u_{i,1}\sqrt{\lambda_1}, u_{i,2}\sqrt{\lambda_2})$ or $(v_{i,1}\sqrt{\lambda_1}, v_{i,2}\sqrt{\lambda_2})$ might cover the null point (0, 0).

The 0.95 HPD regions for the other genotype scores were not drawn (and not shown in Table 1) because they contained the null point (0,0), which evidence that their contribution to the interaction was not statistically significant. Since there is a clear overlapping of the GE interaction scores for genotypes 1 and 2 at the 0.90 HPD and 0.95 HPD regions, we may conclude that among these two genotypes there are not statistical different in the GE interaction effects. On the other hand, since there is no overlapping of the GE interaction scores for genotypes 5 versus genotypes 1 and 2 at any of the 0.90 HPD and 0.95 HPD, we may conclude that among these two groups of genotypes there are a different interaction effects that are statistically significant.

Analogously, from Table 1 and Figure 3a it can be seen that the posterior density for scores of S6 and S11 do not include the null point (0, 0) at the 0.95 HPD thus there is enough posterior evidence that the multiplicative GE environment effect is significant at 0.95 probability level. By considering Figure 1a and Figure 3a simultaneously, the response of genotypes, environments, and the joint response of genotypes and environments can be extracted. Genotype 5 formed a group by itself and genotypes 1 and 2 formed another distinct group which shows differential responses on environment S6. Since genotype 5 points in similar direction to S6 we conclude that they had a positive GE, whereas genotypes 1 and 2 points in opposite in direction to S6 indicating no GE or negative GE.

A hierarchical cluster algorithm with a complete linkage strategy based on the posterior means of Euclidean posterior distances between rows of **U** and **V** was performed as an additional descriptive tool. Their dendrograms are presented in Figure 2a and Figure 4a for genotypes and environments, respectively. Genotype 1, 2, and 5 are the farthest score from (0, 0) and significant for the GE interaction; genotype 5 is clustered in the first the main group of genotypes (Figure 2a) while genotypes 3 and 12 are within the other main group. Concerning the clustering of the

environments, S11 does not clustered with any of the other non-significant environments (Figure 4a).

In summary, the hierarchical analysis found significant GE interaction patterns between genotypes and environments and it provides useful information for the joint response of genotypes and environments. This meta-analysis using the Bayesian hierarchical approach for identifying subsets of homogenous genotypes and environments that cause significant GE is useful to the breeders due to all the existing information (prior information and information from each period of evaluation) is included in the analysis of the super population parameters. By considering the individual analysis of each period of evaluation is possible to find significant effects that were not revealed by the meta-analysis, for this reason is important to consider the individual analysis too

Individual year analyses - response of genotypes and environment

From a plant breeding perspective asses the overall adaptation of genotypes to environmental conditions through several periods have the same importance than the evaluation of the specific adaptation in each period. As in the meta-analysis, when year 1 and year 2 are analyzed separately, it is possible to find genotypes and environments that contribute to the GE interaction too. For example, genotypes 1 and 2 and environments S2, S3, S5, S8 and S11 showed significant interaction when years 1 and 2 are analyzed separately (Table 2 and Table 3, Figure 1b-c and Figure 3b-c); genotypes 1 and 2 and the environment S11 were also significant in the results obtained from the meta-analysis. On the other hand, the biplots from the different periods provides information of specific responses that do not show up to be significant in the meta-analysis, i.e., S2, S3, S5, and S8 in year 1 and S3 and S8 in year 2.

In year 1, genotype {1} formed a single group with significant GE (Table 2 and Figure 1b). Similarly, in year 1 the 0.90 and 0.95 HPD regions of the environments S2, S5 and S8 overlap and form one group of environments with significant GE, while the environment S3 forms a single group. A similar pattern was found in year 2 were the genotypes 1 and 2 are clustered in one main group {1, 2}, and the environments that were significantly different from zero are clustered in two groups: {S8, S11} and {S3} (Table 3 and Figure 3c).

CONCLUSIONS

The results from the bayesian hierarchical analysis presented in this research is appropriate for dealing with differences between year effects because all the existing information (prior information and information from each period of evaluation) was included in the analysis; while the analysis proposed by Perez-Elizalde et al (2011) only consider the prior information for the analysis by separately of the others periods. The individual analysis in the proposed model use as prior information the super population parameters, this helps to stabilize the posterior distributions

Due to the meta-analysis naturally incorporates the existing information from all periods (years) the corresponding posterior distributions has a pooled dispersion parameter that decreases as the number of periods increases; this gives more confidence in the final conclusions as compared with those results obtained in the individual analysis.

In this research, the bayesian inference was proposed for the analysis of the linear-bilinear models by considering the multivariate von Mises-Fisher distribution as prior distribution for the interaction parameters, as was proposed by Pérez-Elizalde et al (2011) who developed the bayesian analysis for the orthonormal matrices **U** and **V** through MCMC method.

Much has been discussed about the variation in the response over the years and some inconsistencies in the results are frequently found from one year to another. As was show, the inclusion of several data sources helps to overcoming this difference in the response due to natural variation, usually this difference is considered as year effect. The robustness arises on two principal reasons, first by considering the prior information, and second by the borrowing of information between periods through the super population parameters.

This Bayesian hierarchical model is useful for breeders in the process of assessing and detecting genotypes and environments that cause interaction and it seems less sensible to change from one year to another. This method is easily extended to others linear-bilinear models just by relaxing some model constraints.

Acknowledgments

The authors would like to thank all national collaborators who carried out the multi-environment maize trials include in this study.

REFERENCES

Cornelius, P.L., J. Crossa, and M.S. Seyedsadr. 1996. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. p. 199-234. *In* M.S. Kang and H.G. Gauch (eds.). Genotype-by-environment interaction. CRC Press, Boca Raton, FL.

Crossa, J., R.-C. Yang, 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.

Crossa, J., S. Perez-Elizalde, D. Jarquin, J. M. Cotes, K. Viele, G. Liu, and P.L. Cornelius. 2011. Bayesian estimation of the Additive Main effects and Multiplicative Interaction (AMMI) model. Crop Science 51:1458-1469.

Gabriel, K.R. 1978. Least squares approximation of matrices by additive and multiplicative models. *Journal of the Royal Statistical Society*, Series B, 40:186-196.

Gauch, H.G. 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44:705-715.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). Bayesian Data Analysis, second edition. London: Chapman and Hall.

Gollob, H.F. 1968. A statistical model which combines features of factor analytic and analysis of variance. *Psychometrika* 33:73-115.

Hoff, P.D. 2009. Simulation of the matrix Bingham-von Mises-Fisher distribution, with applications to multivariate and relational data. *Journal of Computational and Graphical Statistics* 18:438-456.

Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. *J. Agric. Sci*. 103:123-135.

Liu, G. 2001. Bayesian computation for general linear-bilinear models. Ph D. dissertation. Department of Statistics. University of Kentucky. Lexington, KY.

Mandel, J. 1969. The partitioning of interaction in analysis of variance. *Journal of Research of the National Bureau of Standards*, Series B, 73:309-328.

Mandel, J. 1971. A new analysis of variance models for non-additive data. *Technometrics* 13:1-18.

Perez-Elizalde, S., Jarquin, D., J. Crossa, J. 2011. A general Bayesian estimation method of linear-bilinear models applied to plant breeding trials with genotype × environment interaction. *Journal of Agricultural, Biology and Environmental Statistics* (*JABES*) doi: 10.1007/s13253-011-0063-9.

Viele, K., and C. Srinivasan. 2000. Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback-Leibler information. *Journal of Statistical Planning and Inference* 84:201-219.

Table 1. Posterior summary (Mean), standard deviation (SD), quartile ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$), 0.95 HPD intervals computed with 20,000 approximately independent samples simulated from the joint posterior distribution for the singular values (λ_1 and λ_2) and the right and left singular vector elements of genotypes and environments, respectively, whose 0.95 HPD intervals do not contain the null value (0, 0). Data for grain yield measured in tons per hectare for the meta-analysis (years 1 and 2 combined).

						0.95 HPD interval	
Parameter	Mean	SD	*q0.25*	*q0.5*	*q0.75*	Lower	Upper
$u_{1,1}$	0.46	0.12	0.38	0.47	0.54	0.22	0.69
$u_{2,1}$	0.33	0.13	0.25	0.34	0.42	0.07	0.58
$u_{5,1}$	-0.35	0.13	-0.44	-0.35	-0.26	-0.60	-0.10
$v_{6,1}$	-0.33	0.14	-0.43	-0.34	-0.24	-0.60	-0.04
$v_{8,1}$	0.31	0.14	0.21	0.31	0.41	0.03	0.58
$v_{11,1}$	0.55	0.12	0.48	0.56	0.64	0.30	0.79
λ_1	2.73	0.54	2.37	2.73	3.09	1.68	3.79
λ_2	1.23	0.51	0.86	1.22	1.58	0.23	2.20

Table 2. Posterior summary (Mean), standard deviation (SD), quartile ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$), 0.95 HPD intervals computed with 20,000 approximately independent samples simulated from the joint posterior distribution for the inverse of the residual variance (τ), the singular values (λ_1 and λ_{11}) and the right and left singular vector elements of genotypes and environments, respectively, whose 0.95 HPD intervals do not contain the null value (0, 0). Data for grain yield measured in tons per hectare for year 1.

						0.95 HPD Interval	
Parameter	Mean	SD	*q0.25*	*q0.50*	*q0.75*	Lower	Upper
τ	2.52	0.18	2.39	2.51	2.64	2.16	2.87
$u_{1,1}$	0.54	0.12	0.47	0.55	0.62	0.30	0.75
$u_{5,1}$	-0.27	0.14	-0.36	-0.27	-0.18	-0.54	0.00
$v_{2,1}$	0.32	0.14	0.23	0.33	0.42	0.05	0.59
$v_{3,1}$	-0.39	0.15	-0.49	-0.40	-0.30	-0.66	-0.09
$v_{5,1}$	0.36	0.15	0.28	0.38	0.47	0.07	0.64
$v_{8,1}$	0.38	0.13	0.30	0.39	0.47	0.13	0.63
λ_1	2.66	0.51	2.34	2.68	3.01	1.67	3.66
λ_2	1.01	0.54	0.58	0.98	1.39	0.00	1.95

Table 3. Posterior summary (Mean), standard deviation (SD), quartile ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$), 0.95 HPD intervals computed with 20,000 approximately independent samples simulated from the joint posterior distribution for the inverse of the residual variance (τ), the singular values (λ_1 and λ_{11}) and the right and left singular vector elements of genotypes and environments, respectively, whose 0.95 HPD intervals do not contain the null value (0, 0). Data for grain yield measured in tons per hectare for year 2.

						0.95 HPD interval	
Parameter	Mean	SD	*q0.25*	*q0.5*	*q0.75*	Lower	Upper
τ	2.29	0.16	2.17	2.28	2.39	1.97	2.61
$u_{1,1}$	0.36	0.13	0.28	0.37	0.45	0.11	0.61
$u_{2,1}$	0.32	0.15	0.23	0.33	0.42	0.02	0.60
$v_{3,1}$	-0.33	0.14	-0.43	-0.34	-0.24	-0.59	-0.04
$v_{6,1}$	-0.30	0.14	-0.39	-0.30	-0.21	-0.56	-0.02
$v_{8,1}$	0.38	0.14	0.29	0.39	0.47	0.11	0.64
$v_{11,1}$	0.44	0.13	0.36	0.45	0.53	0.18	0.69
λ_1	2.57	0.50	2.25	2.58	2.92	1.58	3.56
λ_2	0.95	0.54	0.53	0.91	1.32	0.00	1.91

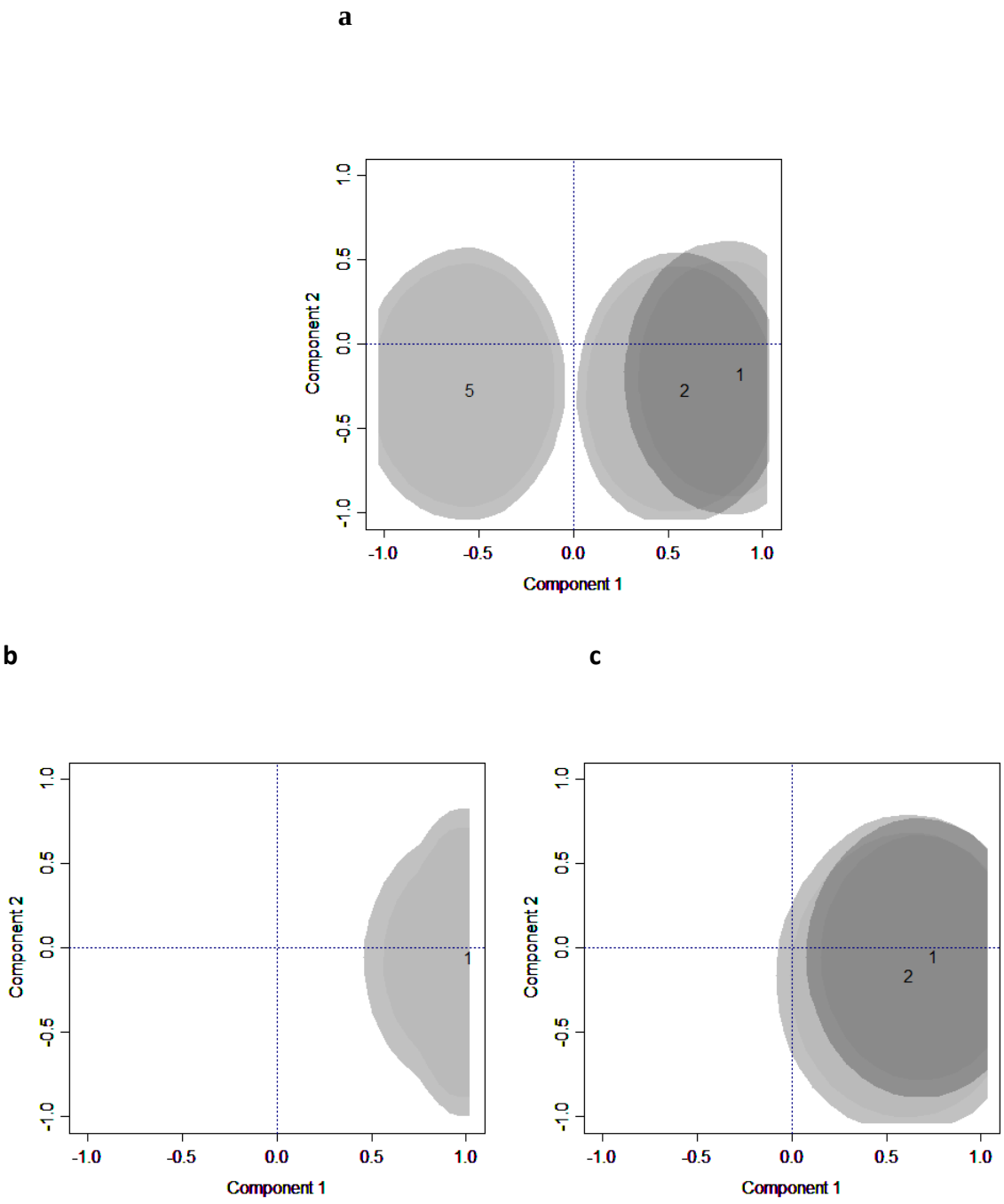


Figure 1. Plant breeding data of 20 genotypes in 12 environments; plot of the row (genotypes) scores $\mathbf{U'D^{1/2}}$ and the bivariate 95% (gray external contour) and 90% (gray internal contour) HPD regions (only the genotypes which do not include the null point (0, 0) at the 95% HPD probability are depicted) for: (**a**) meta-analysis, genotypes (1, 2 and 5) ; (**b**) analysis of year 1, genotype (1); (**c**) analysis of year 2, genotypes (1 and 2).

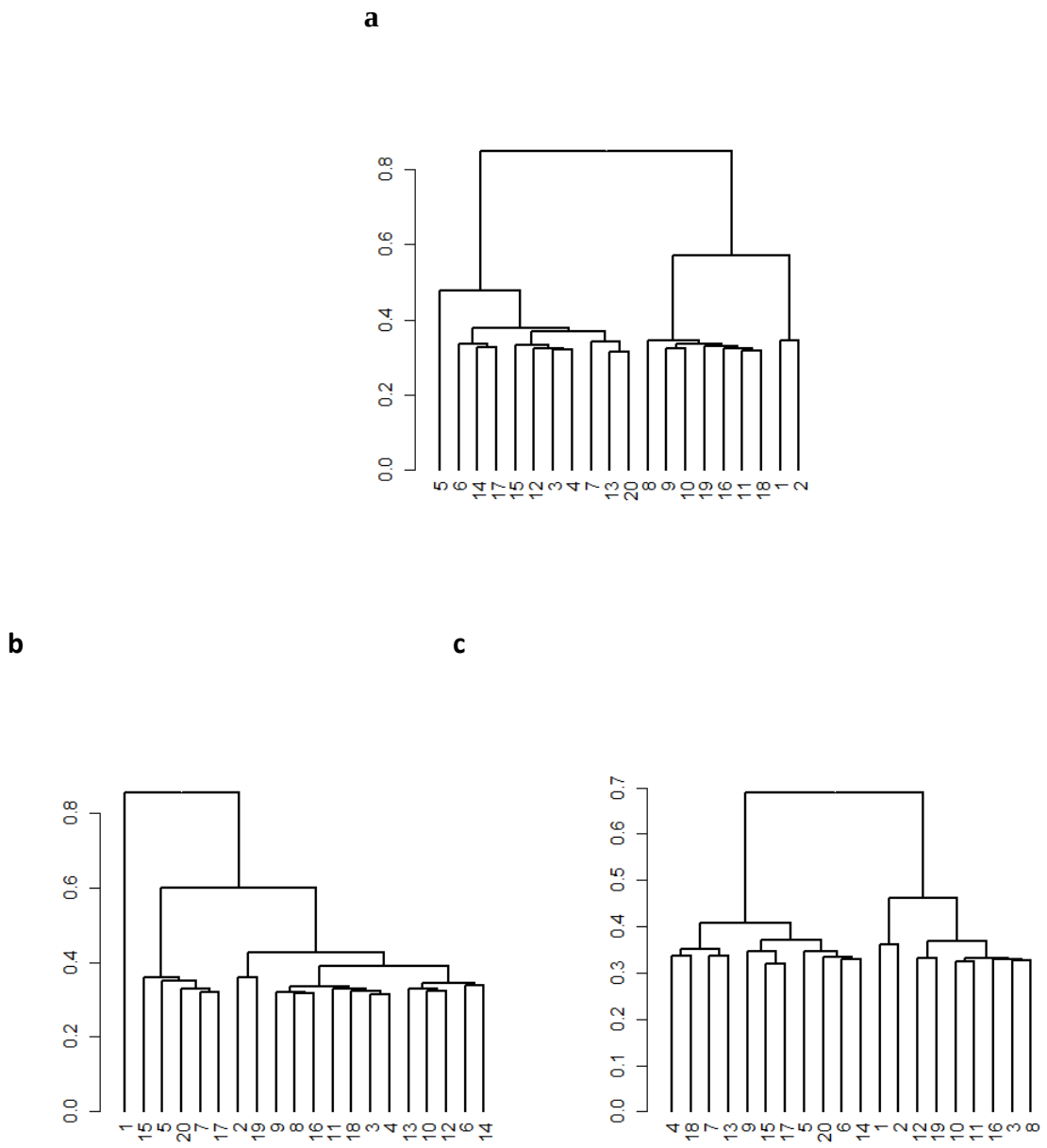


Figure 2. Plant breeding data of 20 genotypes in 12 environments; dendograms of 20 genotypes using the first two eigenvectors for: (**a**) meta-analysis; (**b**) analysis of year 1; (**c**) analysis of year 2.

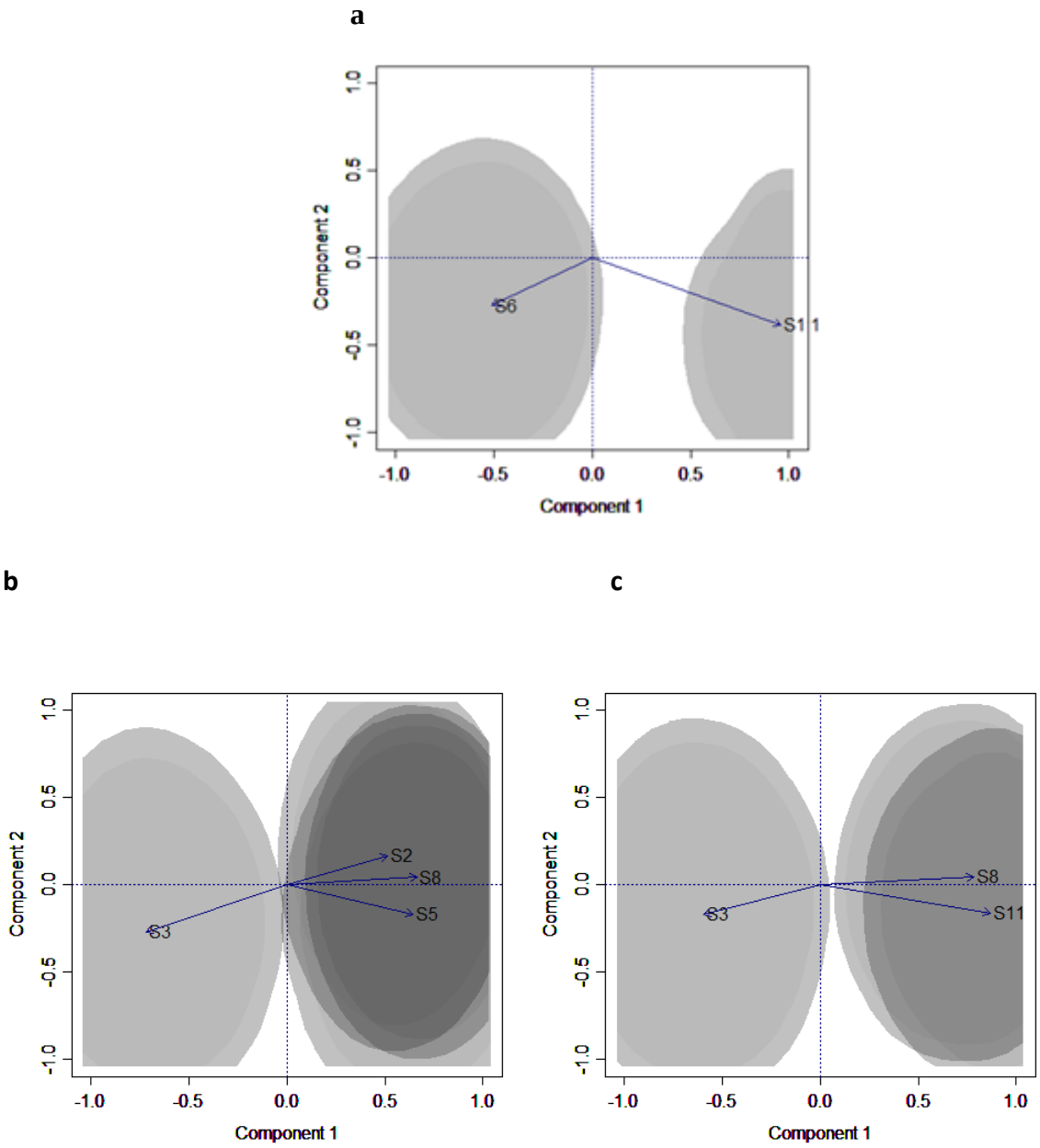


Figure 3. Plant breeding data of 20 genotypes in 12 environments; plot of the column (genotypes) scores $\mathbf{V'D^{1/2}}$ and the bivariate 95% (gray external contour) and 90% (gray internal contour) HPD regions (only the environments which do not include the null point (0, 0) at the 95% HPD probability are depicted) for: (**a**) meta-analysis, environments (S6 and S11) ; (**b**) analysis of year 1, environments (S2, S3, S5 and S8); (**c**) analysis of year 2, environments (S3, S8 and S11).

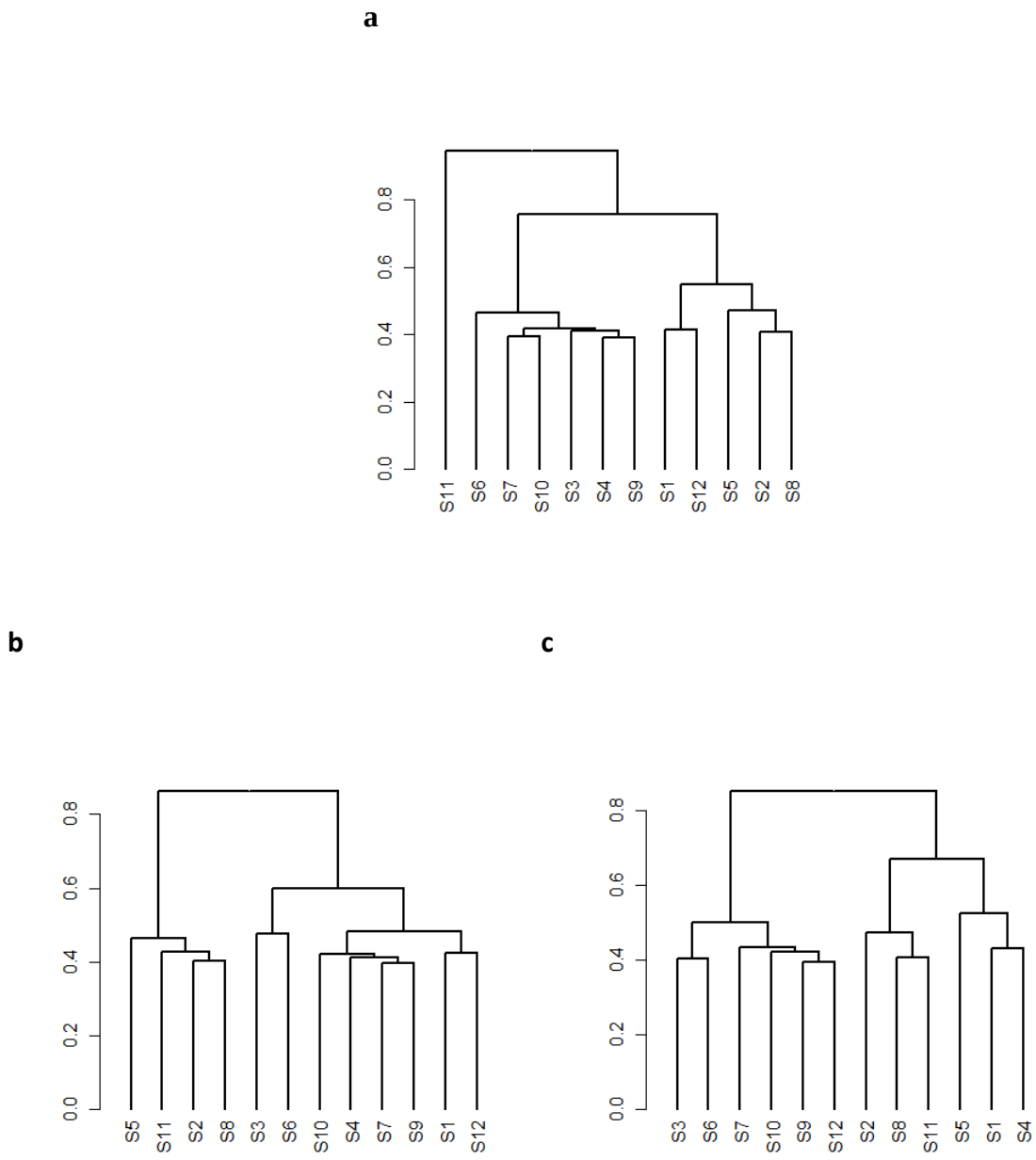


Figure 4. Plant breeding data of 20 genotypes in 20 environments; dendograms of the 12 environments using the first two eigenvectors for: (**a**) meta-analysis; (**b**) analysis of year 1; (**c**) analysis of year 2.

APPENDIX A

Full conditional posterior distributions

$$\pi\left(\boldsymbol{\theta}_m^* \,\middle|\, \tau_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \boldsymbol{\Sigma}_m^*, \mathbf{Y}_m\right) =$$

$$\boldsymbol{N}_{r+c-1}\left(\boldsymbol{\theta}_m^* \,\middle|\, (\textstyle\sum_a^{*-1} + \sum_m^{*-1})^{-1} \left(\sum_a^{*-1} \widehat{\boldsymbol{\theta}}_m^* + \sum_m^{*-1} \boldsymbol{\theta}^*\right), (\sum_a^{*-1} + \sum_m^{*-1})^{-1}\right)$$

where $\widehat{\boldsymbol{\theta}}_m^* = \left(\frac{\mathbf{1}_r^t \mathbf{Y}_m \mathbf{1}_c}{rc}, \frac{\mathbf{K}_r^t \mathbf{Y}_m \mathbf{1}_c}{c}, \frac{\mathbf{K}_c^t \mathbf{Y}_m^t \mathbf{1}_r}{r}\right)$ and $\Sigma_a = (n_m \tau_m)^{-1} \begin{bmatrix} (r*c)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & c^{-1}\mathbf{I}_{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & r^{-1}\mathbf{I}_{c-1} \end{bmatrix}$

$$\pi(\mathbf{D}_m \mid \tau_m, \mathbf{U}_m, \mathbf{V}_m, \mathbf{Y}_m, \mathbf{D}) = \prod_{k=1}^{t} \left\{1 - \Phi\left(\sqrt{(\tau_m(n_m + n_0))^{-1}}\left(\lambda_{(k+1)_m} - \frac{n_m l_{k_m} + n_0 l_k}{n_m + n_0}\right)\right)\right\}^{-1} \times$$

$$\mathrm{N}\left(\lambda_{k_m} \mid \frac{n_m l_{k_m} + n_0 l_k}{n_m + n_0}, (\tau_m(n_m + n_0))^{-1}\right)$$

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

$$\pi(\mathbf{U}_m \mid \tau_m, \mathbf{V}_m, \mathbf{D}_m, \mathbf{Y}_m, \mathbf{M}_{0m}, \mathbf{V}, \mathbf{D}) \propto \mathrm{etr}(\tau_m[n_m \mathbf{Y}_m \mathbf{V}_m \mathbf{D}_m + n_0 \mathbf{M}_{0m} \mathbf{V} \mathbf{D}]\mathbf{U}_m')$$

$$\pi(\mathbf{V}_m \mid \tau_m, \mathbf{U}_m, \mathbf{D}_m, \mathbf{Y}_m, \mathbf{M}_{0m}, \mathbf{U}, \mathbf{D}) \propto \mathrm{etr}(\tau_m[n_m \mathbf{Y}_m' \mathbf{U}_m \mathbf{D}_m + n_0 \mathbf{M}_{0m} \mathbf{U} \mathbf{D}]\mathbf{V}_m')$$

$$\pi(\tau_m \mid a_m, b_m, \mathbf{E}_m) = \mathrm{Ga}(\tau_m \mid \tfrac{n_m rc}{2} + a_m, \tfrac{n_m}{2} \mathrm{tr}(\mathbf{E}_m \mathbf{E}_m') + \frac{(n_m - 1)\mathrm{tr}(\mathbf{S}_m \mathbf{S}_m')}{2} + b_m)$$

$$\pi\left(\boldsymbol{\theta}^* \,\middle|\, \mathbf{U}_m, \mathbf{V}_m, \mathbf{D}_m, \sum_{m=1}^{h} \sum_m^{*-1}, \tau_m, \mathbf{Y}_m\right) =$$

$$N_{r+c-1}\left(\boldsymbol{\theta}^* \,\middle|\, (\sum_{m=1}^{h} \sum_m^{*-1} + \sum_p^{*-1})^{-1} \left(\sum_{m=1}^{h} \sum_m^{*-1} \widehat{\boldsymbol{\theta}}_m^* + \sum_p^{*-1} \boldsymbol{\theta}_0^*\right), (\sum_{m=1}^{h} \sum_m^{*-1} + \sum_p^{*-1})^{-1}\right)$$

$$\pi(\mathbf{D} \mid \tau_m, \mathbf{D}_m) = \prod_{k=1}^{t} \left\{ 1 - \Phi\left(\sqrt{(\sum_{m=1}^{h} n_m \tau_m + \tau_0)^{-1}} \left(\lambda_{k+1} - \frac{\sum_{m=1}^{h} n_m \tau_m l_{k_m} + \tau_0 l_m^0}{\sum_{m=1}^{h} n_m \tau_m + \tau_0} \right) \right) \right\}^{-1}$$

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

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

$$\pi(\mathbf{U} \mid \tau_m, \mathbf{V}_m, \mathbf{M}_{0m}, \mathbf{D}) \propto \mathrm{etr}([n_0 \sum_{m=1}^{h} \tau_m \mathbf{M}_{0m} \mathbf{V}_m \mathbf{D} + \tau_0 \mathbf{M_0} \mathbf{V}_0 \mathbf{D}_0] \mathbf{U}')$$

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

APPENDIX B

Table B1. Posterior summary (Mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$) and 0.95 HPD intervals of the 5000 approximately independent samples simulated from the joint posterior distribution of the 33 linear effects for plant breeding data of grain yield measured in tons per hectare for the meta-analysis.

						0.95 HPD interval	
Parameter	Mean	SD	*q0.25*	*q0.5*	*q0.75*	Lower	Upper
μ	5.00	0.04	4.98	5.00	5.03	4.93	5.08
α_1	-0.01	0.16	-0.12	-0.01	0.10	-0.32	0.29
α_2	-0.11	0.16	-0.21	-0.11	0.00	-0.41	0.20
α_3	-0.12	0.16	-0.22	-0.12	-0.01	-0.43	0.19
α_4	0.00	0.16	-0.11	0.00	0.10	-0.30	0.31
α_5	-0.08	0.16	-0.18	-0.07	0.03	-0.38	0.23
α_6	-0.31	0.16	-0.42	-0.31	-0.21	-0.62	-0.01
α_7	0.11	0.16	0.01	0.11	0.22	-0.19	0.41
α_8	0.21	0.16	0.10	0.21	0.31	-0.09	0.52
α_9	-0.05	0.16	-0.15	-0.05	0.06	-0.37	0.24
α_{10}	0.44	0.15	0.34	0.44	0.55	0.13	0.74
α_{11}	0.04	0.16	-0.06	0.04	0.15	-0.26	0.35
α_{12}	0.06	0.16	-0.04	0.06	0.17	-0.24	0.37
α_{13}	0.27	0.16	0.17	0.27	0.38	-0.03	0.59
α_{14}	-0.21	0.16	-0.31	-0.21	-0.10	-0.52	0.09
α_{15}	0.03	0.16	-0.07	0.03	0.14	-0.28	0.33
α_{16}	-0.23	0.16	-0.34	-0.23	-0.12	-0.54	0.07
α_{17}	-0.06	0.16	-0.16	-0.06	0.05	-0.37	0.24

(Table B1. continued)

α_{18}	-0.04	0.16	-0.15	-0.04	0.06	-0.34	0.27
α_{19}	-0.09	0.16	-0.20	-0.09	0.01	-0.39	0.22
α_{20}	0.13	0.16	0.02	0.13	0.23	-0.18	0.42
β_1	2.29	0.12	2.21	2.29	2.37	2.06	2.52
β_2	2.36	0.12	2.28	2.36	2.44	2.13	2.60
β_3	2.27	0.12	2.19	2.28	2.35	2.04	2.51
β_4	-0.87	0.12	-0.95	-0.87	-0.79	-1.11	-0.64
β_5	-0.51	0.12	-0.59	-0.51	-0.43	-0.75	-0.28
β_6	-0.78	0.12	-0.86	-0.78	-0.70	-1.01	-0.55
β_7	-0.80	0.12	-0.88	-0.80	-0.72	-1.04	-0.56
β_8	-0.16	0.12	-0.24	-0.16	-0.08	-0.39	0.08
β_9	-0.17	0.12	-0.25	-0.17	-0.09	-0.40	0.06
β_{10}	-1.16	0.12	-1.24	-1.16	-1.08	-1.39	-0.92
β_{11}	-0.99	0.12	-1.07	-0.99	-0.91	-1.22	-0.76
β_{12}	-1.49	0.12	-1.57	-1.49	-1.41	-1.72	-1.25

Table B2. Posterior summary (Mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$) and 0.95 HPD intervals of the 5000 approximately independent samples simulated from the joint posterior distribution of the 33 linear effects for plant breeding data of grain yield measured in tons per hectare for the year 1.

						0.95 HPD interval	
Parameter	Mean	SD	*q0.25*	*q0.5*	*q0.75*	Lower	Upper
μ	4.97	0.03	4.96	4.97	4.99	4.92	5.03
α_1	0.02	0.11	-0.05	0.02	0.10	-0.20	0.25
α_2	-0.07	0.11	-0.15	-0.07	0.01	-0.30	0.15
α_3	-0.03	0.12	-0.11	-0.03	0.05	-0.26	0.19
α_4	-0.02	0.11	-0.10	-0.02	0.05	-0.25	0.20
α_5	-0.05	0.12	-0.12	-0.05	0.03	-0.27	0.18
α_6	-0.36	0.12	-0.44	-0.36	-0.28	-0.59	-0.14
α_7	0.07	0.11	0.00	0.07	0.15	-0.16	0.29
α_8	0.14	0.11	0.06	0.14	0.21	-0.09	0.36
α_9	-0.01	0.11	-0.08	-0.01	0.07	-0.23	0.22
α_{10}	0.50	0.11	0.42	0.49	0.57	0.27	0.72
α_{11}	-0.01	0.12	-0.09	-0.01	0.06	-0.24	0.21
α_{12}	0.10	0.12	0.03	0.10	0.18	-0.12	0.34
α_{13}	0.31	0.12	0.24	0.31	0.39	0.09	0.54
α_{14}	-0.19	0.11	-0.27	-0.19	-0.12	-0.41	0.03
α_{15}	0.00	0.11	-0.08	0.00	0.07	-0.22	0.23
α_{16}	-0.24	0.12	-0.31	-0.24	-0.16	-0.46	-0.01
α_{17}	-0.08	0.12	-0.16	-0.08	0.00	-0.31	0.14

(Table B2. continued)

α_{18}	0.03	0.12	-0.04	0.03	0.11	-0.19	0.26
α_{19}	-0.13	0.12	-0.21	-0.13	-0.05	-0.35	0.10
α_{20}	0.02	0.11	-0.06	0.02	0.10	-0.21	0.24
β_1	2.47	0.09	2.41	2.47	2.53	2.29	2.64
β_2	2.65	0.09	2.59	2.65	2.70	2.47	2.82
β_3	2.31	0.09	2.25	2.31	2.37	2.14	2.48
β_4	-0.98	0.09	-1.04	-0.98	-0.92	-1.15	-0.81
β_5	-0.45	0.09	-0.51	-0.45	-0.40	-0.63	-0.29
β_6	-0.94	0.09	-1.00	-0.94	-0.88	-1.11	-0.77
β_7	-0.36	0.09	-0.42	-0.36	-0.30	-0.54	-0.19
β_8	-0.09	0.09	-0.15	-0.09	-0.03	-0.26	0.08
β_9	-0.28	0.09	-0.34	-0.28	-0.22	-0.44	-0.10
β_{10}	-1.36	0.09	-1.42	-1.36	-1.30	-1.53	-1.19
β_{11}	-1.12	0.09	-1.18	-1.12	-1.06	-1.29	-0.95
β_{12}	-1.84	0.09	-1.90	-1.84	-1.79	-2.02	-1.67

Table B3. Posterior summary (Mean), standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$) and 0.95 HPD intervals of the 5000 approximately independent samples simulated from the joint posterior distribution of the 33 linear effects for plant breeding data of grain yield measured in tons per hectare for the year 2.

						0.95 HPD interval	
Parameter	Mean	SD	*q0.25*	*q0.5*	*q0.75*	Lower	Upper
μ	5.04	0.03	5.02	5.04	5.06	4.98	5.09
α_1	-0.05	0.12	-0.13	-0.05	0.03	-0.28	0.19
α_2	-0.15	0.12	-0.23	-0.15	-0.07	-0.39	0.08
α_3	-0.20	0.12	-0.28	-0.20	-0.12	-0.43	0.03
α_4	0.02	0.12	-0.06	0.02	0.10	-0.21	0.26
α_5	-0.11	0.12	-0.19	-0.11	-0.03	-0.34	0.13
α_6	-0.27	0.12	-0.35	-0.27	-0.19	-0.50	-0.03
α_7	0.16	0.12	0.08	0.16	0.24	-0.07	0.39
α_8	0.29	0.12	0.21	0.29	0.37	0.05	0.51
α_9	-0.09	0.12	-0.17	-0.09	-0.01	-0.33	0.14
α_{10}	0.38	0.12	0.30	0.38	0.46	0.15	0.62
α_{11}	0.11	0.12	0.03	0.11	0.19	-0.12	0.34
α_{12}	0.01	0.12	-0.07	0.01	0.09	-0.22	0.24
α_{13}	0.23	0.12	0.15	0.23	0.31	0.00	0.47
α_{14}	-0.22	0.12	-0.30	-0.22	-0.14	-0.46	0.01
α_{15}	0.07	0.12	-0.01	0.07	0.15	-0.16	0.31
α_{16}	-0.23	0.12	-0.31	-0.23	-0.15	-0.46	0.01
α_{17}	-0.02	0.12	-0.10	-0.02	0.06	-0.26	0.21

(Table B3. continued)

α_{18}	-0.13	0.12	-0.21	-0.13	-0.05	-0.36	0.10
α_{19}	-0.05	0.12	-0.13	-0.05	0.04	-0.28	0.19
α_{20}	0.25	0.12	0.17	0.25	0.33	0.02	0.49
β_1	2.09	0.09	2.03	2.09	2.15	1.91	2.26
β_2	2.06	0.09	2.00	2.06	2.12	1.89	2.24
β_3	2.23	0.09	2.17	2.23	2.29	2.05	2.40
β_4	-0.76	0.09	-0.82	-0.76	-0.70	-0.94	-0.57
β_5	-0.56	0.09	-0.62	-0.56	-0.50	-0.74	-0.38
β_6	-0.60	0.09	-0.66	-0.60	-0.54	-0.77	-0.42
β_7	-1.29	0.09	-1.35	-1.29	-1.23	-1.48	-1.11
β_8	-0.23	0.09	-0.29	-0.23	-0.17	-0.41	-0.06
β_9	-0.05	0.09	-0.12	-0.05	0.01	-0.24	0.12
β_{10}	-0.94	0.09	-1.00	-0.94	-0.88	-1.12	-0.76
β_{11}	-0.83	0.09	-0.90	-0.83	-0.77	-1.00	-0.66
β_{12}	-1.10	0.09	-1.16	-1.10	-1.04	-1.27	-0.91

CONCLUSIONES GENERALES

Para el estudio de la interacción, dentro del contexto del análisis de tablas de doble entrada, en este trabajo se planteó una modelación bayesiana de los modelos lineales – bilineales y se propuso como distribución a priori de los parámetros de interacción la distribución von Mises Fisher. El enfoque bayesiano permite realizar inferencia (construcción de regiones credibilidad y pruebas de hipótesis bayesianas) en los parámetros de la interacción y también ofrece la posibilidad de incorporar información disponible. En muchas de las áreas donde se lleva a cabo el estudio de este tipo de interacción generalmente existe información que puede aprovecharse para la obtención de conclusiones más precisas. Las fuentes de información pueden ser provenir del conocimiento del fenómeno por parte de un experto ó de una serie de registros de experimentos realizados con anterioridad. Los alcances y la aplicación de los modelos bayesianos propuestos en esta investigación dependen en gran medida del grado y la cantidad de información disponible. En los diferentes capítulos se desarrollo la teoría para el modelo AMMI el cual ofrece una familia de modelos a partir de relajar algunas restricciones. Los tres modelos fueron ejemplificados con datos de experimentos diseñados para llevar a cabo fitomejoramiento. El modelo del Capítulo 1 puede usarse cuando se dispone de poca o nula información acerca del fenómeno de interés. En el Capítulo 2, el modelo propuesto incorpora información de un experimento realizado con anterioridad mientras que el modelo del Capítulo 3 ofrece la posibilidad de analizar e incorporar información de una serie de experimentos.

LITERATURA CITADA

Asfaw, A.; Alemayehu, F.; Gurum, F.; Atnaf, M. AMMI and GGE Biplot analysis for matching varieties onto soybean production environments in Ethiopia. Scientific Research and Essay, v.4, p.1322-1330, 2009.

Ceccarelli, S.; Grando, S.; Booth, R.H. 1996. International breeding programmes and resource-poor farmers: crop improvements in difficult environments. In "Participatory Plant Brreding" (P. Eyzaguirre and M. Iwanaga, eds.), pp. 99-116. IPGRI, Rome, Italy.

Ebdon, JS, Gauch HG (2002). Additive main effect and multiplicative interaction analysis of national turfgrass performance trials II: Cultivar recommendations. Crop Sci. 42: 497-506.

Gauch HG (2006). Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 46: 1488-1500.

Hoff, P. D. (2009), "Simulation of the Matrix Bingham-von Mises-Fisher Distribution, With Aplications to Multivariate and Relational Data," Journal of Computational and Graphical Statistics, 18, 43-456.

Kaya, Y., M. Akcura, and S. Taner. 2006. GGE-Biplot Analysis of Multi-Environment Yield Trials in Bread Wheat. Turk. J. Agric. For. 30:325-337.

Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. J. Agric. Sici. (Cambridge) 103: 123-135.

Mardia, K. V., J.T., Kent, and J.B. Bibbi. 1979 Multivariate analysis. Academic Press, London, UK.

Matus-Cadiz, M.A., P. Hulcl, C.E.Perron and R. T. Tyler. 2003. Genotype x environment interaction for grain color in hard white spring wheat. Crop Sci. 43: 219-226.

Navabi A, Yang RC, Helm J, Spawer DM (2006). Can spring wheat-growing megaenvironments in the northern great palin be dissected for representative locations or niche-adapted genotypes? Crop Sci. 46:1107-1116.

Samonte, S.O.P.B., L.T. Wilson, A.M. McClung and J.C. Medley. 2005. Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analisys. Crop Sci. 45: 2414-2424.

Yan W, Kang MS, Ma B, Wood S, Cornelius PL (2007). GGE biplot vs. AMMI analysis of genotype-by-enviroment data. Crop Sci. 47: 643-655.

Yan W, Tinker NA (2006). Biplot analysis of multi-environment trial data: Principales and applications. Can. J. Plant Sci. 86: 623-645.

Yan, W., L. A. Hunt, Q. Sheng and Z. Szlavnics. 2000. Cultivar evaluation and mega-environment investigation based on the GEE Biplot. Crop Sci. 40: 597-605.

Viele , K., and C. Srinivasan. 2000 Parsimonius estimation of multiplicative interaction in analysis of variance using Kull-bach-Leibler information. J. Stat. Plan. Infer. 84:201-219.