



CARLOS PEREIRA DA SILVA

**USO DO ALGORITMO *REVERSIBLE JUMP MARKOV*
CHAIN MONTE CARLO PARA SELEÇÃO DE TERMOS
MULTIPLICATIVOS DO MODELO AMMI**

LAVRAS - MG

2019

Carlos Pereira da Silva

**USO DO ALGORITMO *REVERSIBLE JUMP MARKOV CHAIN MONTE*
CARLO PARA SELEÇÃO DE TERMOS MULTIPLICATIVOS DO
MODELO AMMI**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

Orientador

Prof. Dr. Marcio Balestre

LAVRAS - MG

2019

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Silva, Carlos Pereira da.

Uso do algoritmo *Reversible Jump Markov Chain Monte Carlo* para seleção de termos multiplicativos do modelo AMMI / Carlos Pereira da Silva. – Lavras: UFLA, 2019.

128 p. : il.

Tese (doutorado) - Universidade Federal de Lavras, 2019.

Orientador: Marcio Balestre.

Bibliografia.

1. Reversible Jump. 2. AMMI-Bayesiano. 3. Análise Multiambiental. I. Balestre, Marcio. II. Título.

CARLOS PEREIRA DA SILVA

**USO DO ALGORITMO *REVERSIBLE JUMP MARKOV CHAIN MONTE*
CARLO PARA SELEÇÃO DE TERMOS MULTIPLICATIVOS DO
MODELO AMMI**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

APROVADO em 12 de abril de 2019.

Dr. Júlio Sílvio de Sousa Bueno Filho
Dr. Augusto Ramalho de Moraes
Dr. Helvécio Geovani Fagnoli Filho
Dr. Paulo Eduardo Teodoro

UFLA
UFLA
UFLA
UFMS

Prof. Dr. Marcio Balestre
Orientador

**LAVRAS - MG
2019**

À minha mãe pela força, carinho e dedicação, ao longo dos anos.

DEDICO

AGRADECIMENTOS

A Deus, razão da minha força para prosseguir até o fim, e pelas pessoas maravilhosas que fizeram ou fazem ainda parte na minha caminhada da vida.

À minha mãe, Josefa Barbosa de Almeida, pelo amor, dedicação e incentivo para continuar os estudos.

Ao meu irmão Ronaldo e minha cunhada Lívia por incentivar a continuar seguindo.

Aos colegas de doutorado em Estatística e Experimentação Agropecuária, em especial, Cristian Tiago, Fernando Ribeiro, Joel Nuvunga, Sérgio Domingos e Ernandes Guedes. Em especial ao grande amigo Luciano Antônio de Oliveira e companheiro durante a batalha.

À minha noiva Andrezza Kellen Alves Pamplona pelo carinho, compreensão e ajuda.

Aos professores do departamento, em especial Júlio Sílvio de Sousa Bueno Filho, pelas contribuições e pela disposição em ajudar.

Ao professor Dr. Marcio Balestre, pela orientação, confiança, incentivo e atenção durante todas as etapas deste projeto.

Aos funcionários do Departamento de Estatística, em especial à secretária de Pós-Graduação, Nádia Ferreira.

À Universidade Federal de Lavras (UFLA) e ao Departamento de Estatística (DES) pela oportunidade concedida para realização do doutorado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) Código de Financiamento 001 e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes

(Martin Luther King)

RESUMO

O modelo de efeitos principais aditivos e interação multiplicativa (AMMI) adquiriu grande aplicabilidade para análise de dados provenientes de ensaios multiambientais. A determinação de quantos termos bilineares são necessários para explicar a interação genótipos por ambientes (GEI) tem sido exaustivamente estudada no contexto da análise AMMI frequentista. No contexto bayesiano (AMMI-Bayesiano), esse problema tem sido abordado por meio de critérios de informação e fator de Bayes. Contudo, esses procedimentos envolvem computação intensiva inviabilizando aplicações quando o espaço do modelo é grande. O principal objetivo deste trabalho é propor a determinação do número de termos multiplicativos em modelo AMMI-Bayesiano utilizando o algoritmo *Reversible Jump*. Para tanto foram consideradas três versões do AMMI-bayesiano, que se diferenciam apenas pela priori assumida para o parâmetro de escala dos valores singulares, são elas: BAMMI fundamentado no princípio da razão insuficiente (uniforme); BMMIS que tem por base o princípio da invariância (priori de Jeffreys) e ainda BMMIE que utiliza o conceito de priori de máxima entropia. Para exemplificar o método foi simulado um conjunto de dados em que 20 genótipos foram testados em nove ambientes em delineamento de blocos casualizados, com três repetições, cuja variável foi a produtividade. A avaliação preditiva indicou que a análise AMMI, de modo geral, é robusta e o *Reversible Jump* se mostrou um bom método para ajuste e seleção de modelos, sendo que a correlação entre valores observados e preditos por esse método foi sempre maior em comparação ao AMMI ajustado por meio de cadeias MCMC (*Markov Chain Monte Carlo*) convencionais. Biplots condicionais aos modelos mais prováveis e, ainda, marginais em relação aos termos bilineares foram implementados e apresentaram praticamente o mesmo padrão. A questão da seleção de modelos em análise AMMI na perspectiva Bayesiana não tem sido muito abordada na literatura atual e nesse sentido, estudos, como o aqui proposto, são de grande importância para incentivar e viabilizar a utilização do método AMMI-Bayesiano que, mesmo sendo bem fundamentado, ainda não se tornou um procedimento comum de análise multiambiental.

Palavras-chave: Análise Bayesiana. Estabilidade. Interação genótipos por ambientes. Seleção de componentes.

ABSTRACT

Additive Main effects and Multiplicative Interaction model (AMMI) has acquired great applicability for the analysis of data from multi-environmental trials. The determination of how many bilinear terms are required to explain the genotype environment interaction (GEI) has been exhaustively studied in the context of the AMMI frequentist analysis. In the Bayesian context (AMMI-Bayesian), this problem has been approached through information criteria and Bayesian factor. However, these procedures involve intensive computation, making applications impossible when model space is large. The main objective of this work is to propose the determination of the number of multiplicative terms in AMMI-Bayesian model using the Reversible Jump algorithm. For that, three versions of the AMMI-Bayesian were considered, which differ only by the assumed priori for the scale parameter of the singular values: BAMMI based on the principle of insufficient reason (uniform); BAMMIS based on the invariance principle (Jeffreys priori) and BAMMIE, which uses the concept of priori of maximum entropy. To exemplify the method, a data set was simulated in which 20 genotypes were tested in nine environments in a randomized block design, with three replicates, whose variable was productivity. The predictive evaluation indicated that AMMI analysis, in general, is robust and Reversible Jump proved to be a good method for adjustment and selection of models, and the correlation between values observed and predicted by this method was always greater in comparison to AMMI adjusted by conventional Markov chain Monte Carlo (MCMC). Biplots conditional on the most probable models, and also marginal in relation to the bilinear terms, were implemented and presented practically the same pattern. The question of the selection of models in AMMI analysis in the Bayesian perspective has not been much approached in the current literature and, in this sense, studies, as proposed here, are of great importance to encourage and to make feasible the use of AMMI-Bayesian method that, even being well-founded, has not yet become a common multi-environmental analysis procedure.

Keywords: Bayesian analysis. Stability. Genotype environment interaction. Selection of components.

LISTA DE FIGURAS

Figura – 1	Traço e densidades das cadeias para os comentos do modelo BAMMI-RJ.	78
Figura – 2	Valores dos critérios de informação AIC, AICM e BIC, bem como para log-verossimilhança (Log-Lik), Erro quadrático médio (EQM) e variância residual (σ_e^2) em função da dimensão do modelo, para família de modelos BAMMI, BAMMIS e BMMIE.	79
Figura – 3	Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMI-8, cujas regiões não incluem a origem (0,0) usando algoritmo Gibbs.	83
Figura – 4	Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMIS-8, cujas regiões não incluem a origem (0,0) usando algoritmo Gibbs.	84
Figura – 5	Regiões de credibilidades para os escores genotípicos e ambientais do modelo BMMIE-2, cujas regiões não incluem a origem (0,0) usando algoritmo Gibbs.	85
Figura – 6	Probabilidade a posteriori para dimensão (t) do modelo. . .	86
Figura – 7	Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional ao modelo mais provável pelo Reversible Jump (cor azul) e para o modelo marginal (cor vermelha): a) BAMMI3; b) BAMMIS3; c) BMMIE2.	90

Figura – 8	Regiões bivariadas a 95% de credibilidades para o modelo BAMMI, construídas com o modelos marginal (a) e condicional (BAMMI3) (b).	91
Figura – 9	Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMI cujas regiões não contêm a origem (0,0) (modelo marginal).	92
Figura – 10	Regiões de credibilidades para o modelo BMMIS, construídas com base na média marginal a posteriori (a) e média condicional a posteriori (b).	93
Figura – 11	Regiões de credibilidades para os escores genotípicos e ambientais do modelo BMMIS cujas regiões não contêm a origem (0,0).	94
Figura – 12	Regiões de credibilidades para o modelo BMMIE, construídas com base na média marginal a posteriori (a) e média condicional a posteriori (b).	95
Figura – 13	Regiões de credibilidades para os escores genotípicos e ambientais do modelo BMMIE cujas regiões não contêm a origem (0,0).	96
Figura – 14	Avaliação média da capacidade dos modelos com base na Correlação entre o valor observado e o estimado, utilizando o algoritmo Gibbs.	97
Figura – 15	Avaliação média da capacidade dos modelos com base na PRESS, utilizando o algoritmo Gibbs.	98
Figura – 16	Avaliação da capacidade dos modelos com base na Correlação entre o valor observado e o estimado, utilizando o algoritmo <i>Reversible Jump</i>	99

Figura – 17 Avaliação da capacidade dos modelos com base na PRESS, utilizando o algoritmo Reversible Jump.	100
Figura – 18 Efeito genotípico para os modelos BAMMIE2 (AICM - Vermelho), BAMMIS3 (BIC - Azul) e BAMMI4 (AIC - Verde), ajustados com algoritmo de Gibbs.	125
Figura – 19 Efeito genotípico para os modelos BAMMIS2 (AICM - Vermelho) e BAMMIS3 (AIC/BIC - Azul), ajustados com algoritmo de Gibbs.	126
Figura – 20 Efeito genotípico para o modelo BAMMI2 com base no critério BIC/AIC/AICM, ajustados com algoritmo de Gibbs.	126
Figura – 21 Traço e densidades das cadeias para os comentes do modelo BAMMIS-RJ.	127
Figura – 22 Traço e densidades das cadeias para os comentes do modelo BAMMIE-RJ.	127

LISTA DE TABELAS

Tabela – 1	Valores singulares estimados para os modelos AMMI conforme o número de eixos incorporados ao modelo, e sua soma de quadrado da interação acumulada (Algoritmo Gibbs).	80
Tabela – 2	Valores singulares obtidos por meio de decomposição de valores singulares da matriz de interação simulada, assim como o percentual referente a cada eixo principal como o acumulado em cada.	81
Tabela – 3	Valores singulares estimados para o modelo AMMI referentes às três hipóteses realizadas sob a interação, utilizando o algoritmo <i>Reversible Jump</i> .	87
Tabela – 4	Médias a posteriori, desvio padrão (Dp) e limites inferiores (LI) e superiores (LS) de regiões a 95% de credibilidade das variâncias (Var.) genotípicas (σ_g^2) e residuais (σ_e^2) para os modelos condicionais (mais prováveis pelo método <i>Reversible Jump</i>) e modelos marginais referente as três versões AMMI bayesiano analisadas.	88
Tabela – 5	Estimativas pontuais para a componente de variância do erro e da variância do efeito genotípico, referente ao modelo BAMMI, para as possíveis dimensões, utilizando algoritmo Gibbs.	119
Tabela – 6	Estimativas pontuais para a componente de variância do erro e da variância do efeito genotípico, referente ao modelo BAMMIS, para as possíveis dimensões, utilizando algoritmo Gibbs.	119

Tabela – 7	Estimativas pontuais para a componente de variância do erro e da variância do efeito genotípico, referente ao modelo BMMIE, para as possíveis dimensões, utilizando algoritmo Gibbs.	120
Tabela – 8	Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional (BMMI3) e marginal para o modelo BMMI.	120
Tabela – 9	Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional (BMMIS3) e marginal para o modelo BMMIS.	121
Tabela – 10	Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional (BMMIE2) e marginal para o modelo BMMIE.	122
Tabela – 11	Correlação e Press para as diferentes suposições feitas sob o modelo AMMI (Reversible Jump).	122
Tabela – 12	Classificação do padrão dos escores genotípicos pela representação biplot AMMI2 com base no cenário simulado (algoritmo Gibbs padrão).	123
Tabela – 13	Classificação do padrão dos escores genotípicos pela representação biplot AMMI2 com base no cenário simulado (algoritmo RJMCMC).	123
Tabela – 14	Tempo computacional total, em minutos (mins), para o processo de amostragem das três versões AMMI-Bayesiano considerando ajuste por critérios de informação e pelo método RJMCMC.	124

SUMÁRIO

1	INTRODUÇÃO	17
2	REFERENCIAL TEÓRICO	22
2.1	Interação genótipos por ambientes	22
2.2	Métodos para análise de dados multiambientais	23
2.3	O modelo AMMI	24
2.4	Determinação do número de termos bilineares no modelo AMMI	27
2.5	Inferência Bayesiana	28
2.6	Incorporação de informações a priori	29
2.7	Estimação dos parâmetros	30
2.8	Cadeias de Markov e Inferência via técnicas MCMC	32
2.8.1	O algoritmo Metropolis-Hastings	34
2.8.2	Amostrador de Gibbs	35
2.9	O modelo AMMI-bayesiano	36
2.10	Critério de seleção de modelos em inferência bayesiana	38
2.10.1	Critério de Informação de Akaike (AIC)	39
2.10.2	Critério de Informação Bayesiano (BIC)	40
2.10.3	Critério de Informação de Akaike Monte Carlo (AICM)	40
2.11	<i>Reversible Jump</i> em cadeias de Markov	41
2.12	A distribuição invariante	42
2.13	Gerando a proposta para um novo estado da cadeia de Markov	43
2.13.1	Condição de reversibilidade	44
2.13.2	Probabilidade de aceitação	47

2.13.3	Proposta determinística	50
2.13.4	Mapeamento Identidade	52
3	MATERIAL E MÉTODOS	55
3.1	Material	55
3.2	Métodos	55
3.3	Especificação das distribuições a priori para os parâmetros do modelo	56
3.4	Distribuições condicionais completas a posteriori para os parâmetros do modelo	59
3.5	Processo de amostragem via MCMC com <i>Reversible Jump</i> . . .	68
3.5.1	Regra de decisão	68
3.5.2	Algoritmo Gibbs	70
3.5.3	Algoritmo RJMCMC	73
3.6	Capacidade preditiva do modelo	75
3.7	Diagnóstico das cadeias de Markov e inferência a posteriori . .	76
4	RESULTADOS	77
4.1	Convergência e efeitos principais genótipos	77
4.2	Seleção de modelo utilizando critério de informação e efeitos principais de genótipos	78
4.2.1	Regiões de Credibilidades para os modelos AMMI-Bayesiano .	82
4.3	Ajuste do modelo AMMI pelo método <i>Reversible Jump</i>	85
4.3.1	Regiões de credibilidades bivariadas em <i>Reversible Jump</i>	90
4.3.1.1	Modelo BAMMI-RJ	90
4.3.1.2	Modelo BMMIS-RJ	93
4.3.1.3	Modelo BMMIE-RJ	94
4.4	Avaliação da capacidade preditiva dos modelos	96

5	DISCUSSÃO	101
6	CONCLUSÃO	106
	REFERÊNCIAS	107
	APÊNDICES A	115
	APÊNDICES B	119
	APÊNDICES C	125

1 INTRODUÇÃO

Em programas de melhoramento de plantas, a seleção e a recomendação de genótipos que proporcionem boas características de interesse comercial assumem grande importância. O desempenho observado (fenótipo) é produto da composição genética da planta (genótipo), das condições em que a planta se desenvolve (ambiente) e, ainda, da interação entre o genótipo e o ambiente (*Genotype by Environment Interaction - GEI*) (CARVALHO et al., 2002; MALOSETTI; RIBAUT; EEUWIJK, 2013).

A GEI reduz a previsibilidade do desempenho de genótipos em ambientes de interesse, dificultando o trabalho do melhorista na recomendação ampla de genótipos. Essa interação pode ocorrer para características quantitativas de importância econômica e é frequentemente estudada em melhoramento de plantas para reprodutibilidade, longevidade, altura, peso, rendimento e resistência a doenças (DIA; WEHNER; ARELLANO, 2016). Nesse contexto, a modelagem eficiente da GEI tem se tornado um aspecto central em análises estatísticas de conjuntos de dados multiambientais (*Multi-Environmental Trials - MET*). O interesse, então, é desenvolver métodos estatísticos que visem à identificação de genótipos estáveis (que não contribuem significativamente com a interação) e combinações satisfatórias entre genótipos e ambientes a fim de explorar o efeito positivo da GEI.

Diversos métodos foram propostos para a análise de dados METs, tanto univariados quanto multivariados, paramétricos e não paramétricos. Dentre eles, destaca-se o uso de Modelos lineares-bilineares (CORNELIUS; SEYEDSADR, 1997), que combinam análise univariada e multivariada em uma única abordagem, permitindo separar ruído de padrão. No contexto do melhoramento de plantas, dois desses modelos têm sido amplamente utilizados: o modelo de efeitos principais aditivos e interação multiplicativa (*Additive Main Effects and Multiplicative*

Interaction - AMMI) proposto por Gauch (1988), que, como o próprio nome sugere, separa os efeitos de genótipos e de ambientes do efeito da interação entre os fatores (GEI); e o modelo de efeitos principais de genótipos mais interação (*Genotype Main Effect Plus Genotype-by-Environment Interaction - GGE*) (YAN et al., 2000), em que os efeitos de genótipo e da interação são confundidos. Uma variante do modelo GGE, com ampla aplicação na análise de dados multiambientes, foi proposta por Yan et al. (2000) e considera, de forma arbitrária, apenas os dois primeiros componentes bilineares na representação gráfica biplot - GGEbiplot (dois primeiros eixos da decomposição por valores singulares da matriz de GGE).

Existe uma vasta literatura abordando vantagens e limitações desses métodos (BALESTRE et al., 2009; GAUCH, 2006; YAN; GLOVER; KANG, 2010). Os modelos AMMI e GGE, dentre os métodos de análise frequentista (efeitos fixos), oferecem vantagens sobre métodos mais tradicionais (por exemplos, aqueles baseados em regressão simples de genótipos em ambientes). Tais modelos permitem, em uma única análise, estudar adaptabilidade e estabilidade, além de proporcionar a representação gráfica, total ou parcialmente, da interação em poucos eixos principais - sendo esse o padrão de GEI (CROSSA et al., 2011; FINLAY; WILKINSON, 1963; KEMPTON, 1984; GAUCH; PIEPHO; ANNICCHIARICO, 2008).

Vários estudos têm sido conduzidos com esses modelos e um dos principais aspectos discutidos, sobretudo na análise AMMI, é a determinação do número de termos multiplicativos que devem ser retidos para descrever o padrão da interação. Dentre as propostas, para esse fim, destacam-se os métodos utilizando testes *F* aproximados (CORNELIUS; SEYEDSADR; CROSSA, 1992) e os baseados em reamostragem por meio de computação intensiva (CORNELIUS; CROSSA, 1999; DIAS; KRZANOWSKI, 2003; GAUCH; PIEPHO; ANNICCHI-

ARICO, 2008) que se tornaram procedimentos corriqueiros em análise AMMI. Apesar da intensa utilização da análise de variância para estudar a GEI, deve-se ressaltar que não se pode garantir que os componentes principais tenham, de fato, uma distribuição qui-quadrado, sendo essa uma crítica na aplicação dos testes F para determinação do número de termos multiplicativos.

Uma forma de contornar essa dificuldade também é proposta por Cornelius et al. (1996) e Cornelius e Crossa (1999), que utilizam estimadores com efeito de encolhimento para os parâmetros bilineares em modelo AMMI. Esse método dispensa o uso de testes para seleção de termos bilineares e, em geral, modelos assim ajustados se mostraram superiores àqueles em que o número de termos multiplicativos foi determinado a partir de testes estatísticos. Contudo, o procedimento para obter as predições dos valores singulares frequentemente viola as restrições do modelo, sendo então necessário incorporar uma correção, sob essa restrição.

As limitações presentes na análise AMMI convencional (que considera parâmetros com efeitos fixos) motivaram a procura por abordagens mais flexíveis e eficientes como, por exemplo, o uso de modelos mistos via análise de fatores, proposto por Piepho (1998) e Smith, Cullis e Gilmour (2001). Outra abordagem é a condução da análise AMMI sob o enfoque Bayesiano, apresentada por Vilele e Srinivasan (2000) e Liu (2001), que mostraram como conduzir, de forma correta, o processo de amostragem, sobretudo para os parâmetros que descrevem a interação, cujo suporte para a densidade conjunta a posteriori não é trivial. A flexibilidade do AMMI-Bayesiano em incorporar inferência aos termos multiplicativos no biplot, efeitos aleatórios para genótipos e informações anteriores a experimentos, bem como permitir a seleção do número de termos multiplicativos sem violar as restrições do modelo foi ilustrada em vários trabalhos (CROSSA et al., 2011; PEREZ-ELIZALDE; JARQUIN; CROSSA, 2012; OLIVEIRA et al., 2015; SILVA

et al., 2015).

No contexto Bayesiano, a seleção de modelos em análise AMMI tem sido realizada por meio de critérios de informação, tais como Critério de Informação de Akaike (AIC); Critério de Informação Bayesiano (BIC) e Fator de Bayes (FB) (LIU, 2001; PEREZ-ELIZALDE; JARQUIN; CROSSA, 2012). Silva et al. (2015), por sua vez, apresentaram uma justificativa bayesiana para o método Shrinkage (CORNELIUS; CROSSA, 1999), assumindo priori informativa para os valores singulares, e mostraram que o modelo assim ajustado foi superior em predição àqueles ajustados por critérios de informação ou, ainda, pelo fator de Bayes, quando se atribui priori constante para valores singulares, como apresentado por Crossa et al. (2011), Perez-Elizalde, Jarquin e Crossa (2012) e Oliveira et al. (2015). Tal abordagem pode oferecer ajuste de modelos mais parcimoniosos, caso algum procedimento seja incorporado para determinar o número de eixos a serem retidos, pois espera-se a determinação de poucos valores singulares significativos e uma justificativa mais plausível para o uso da representação biplot AMMI-2, já que os primeiros eixos são preditos de forma mais acurada.

Em muitas aplicações, é razoável assumir que existe incerteza também em relação ao modelo que melhor se ajusta a um conjunto de dados. Do ponto de vista Bayesiano, esta incerteza pode ser simplesmente incorporada ao problema de inferência, considerando, o próprio modelo como mais um parâmetro desconhecido a ser estimado e, com isso, os diferentes modelos têm uma distribuição de probabilidade.

Neste sentido, um método que vem ganhando notoriedade em incorporar essa incerteza é o de Saltos Reversíveis em Monte Carlo via Cadeias de Markov (*Reversible Jump Markov Chain Monte Carlo - RJMCMC*), apresentado em Green (1995). A razão para utilizá-lo é conceitual, por ser uma generalização da teoria

de Cadeias de Markov. Com isso, seria interessante assumir que o número de eixos principais a serem mantidos no modelo AMMI é uma variável aleatória. Além disso, entre as abordagens transdimensionais propostas nos últimos anos, o método do *RJMCMC* parece ser o mais efetivo, como argumenta Sisson (2005).

Sendo assim, o principal objetivo deste trabalho é o ajuste do modelo AMMI sob a perspectiva Bayesiana utilizando a metodologia de *RJMCMC*, de modo que a inferência conjunta sobre o número de termos bilineares responsáveis pelo padrão da interação GEI e as regiões de credibilidade correspondentes seja realizada de forma direta no processo amostral, pela inclusão ou não de termos multiplicativos no modelo. Além disso, figuram-se como objetivos específicos: i) avaliar a sensibilidade do modelo AMMI-Bayesiano a especificação de prioris para a componente da variância associada aos valores singulares; ii) utilizar critérios de informação para seleção de modelos, bem como avaliar tais critérios sob as diferentes prioris especificadas, sob simulação e iii) avaliar a capacidade preditiva do modelo AMMI-Bayesiano dentro de um cenário simulado de perdas aleatórias de genótipos em ambientes.

2 REFERENCIAL TEÓRICO

2.1 Interação genótipos por ambientes

O estudo da interação entre genótipos e ambientes (GEI) assume importância fundamental, influenciando significativamente nas conclusões e dificultando o trabalho do melhorista em identificar genótipos superiores. No que se refere à produtividade, por exemplo, as análises buscam identificar genótipos com maior produtividade, estabilidade de produção e ampla adaptabilidade aos mais variados ambientes (SILVA; DUARTE, 2006).

A GEI é percebida pelas respostas diferenciadas de genótipos em função dos ambientes de teste (ano, local, época de semeadura, entre outros) e configura-se no principal motivo em se conduzir experimentos em vários ambientes, de forma a isolar seu efeito (CROSSA, 1990; CROSSA; YANG; CORNELIUS, 2004; YAN; KANG, 2002).

De forma geral, a GEI é classificada como simples quando a resposta dos genótipos em relação aos diferentes ambientes ocorre sem alteração em suas posições relativas. Por outro lado, se ocorrem alterações na classificação de genótipos em diferentes ambientes, a interação é dita complexa, sendo a mais importante a ser considerada na análise (CRUZ; CARNEIRO, 2006; YAN; KANG, 2002).

A presença de interações complexas e significativas em conjuntos de dados MET tem motivado a procura por métodos estatísticos mais flexíveis que sejam capazes de captar e analisar, eficientemente, essa fonte de variação. O estudo da interação permite avaliar a adaptabilidade e a estabilidade de genótipos, tendo por base seu rendimento ou outra característica de interesse; além disso, a utilização de métodos eficazes permite a exploração dos efeitos positivos dessa interação (DUARTE; VENCOVSKY, 1999).

2.2 Métodos para análise de dados multiambientais

Historicamente, a GEI tem sido modelada por diversos métodos estatísticos. Contudo, os métodos multivariados têm sido os mais indicados, pois permitem explorar de forma mais eficaz as informações presentes nos dados. Dentre os procedimentos multivariados pode-se destacar: análise discriminante, análise de agrupamento, análise de componentes principais e os modelos de efeitos principais aditivos (AMMI) e de efeitos principais de genótipo mais interação (GGE) (CROSSA, 1990).

Para os melhoristas de plantas, modelos lineares-bilineares, como os de efeitos fixos GGE e AMMI, oferecem uma maneira mais eficiente de modelar os efeitos de G+GEI ou GEI, respectivamente, em relação à simples regressão de genótipos em ambientes como proposto por Finlay e Wilkinson (1963), Eberhart e Russell (1966). Obviamente, como qualquer abordagem em que os efeitos são tidos como fixos, os modelos lineares-bilineares convencionais estão sujeitos a limitações, tais como dificuldades em lidar com dados desbalanceados e heterogeneidade de variância, o que motivou a elaboração de métodos mais flexíveis.

Abordagens com modelos lineares-bilineares para dados METs considerando efeitos aleatórios foram inicialmente apresentadas em Piepho (1997, 1998), Smith, Cullis e Gilmour (2001) e Piepho e Möhring (2006), sendo propostas para os modelos AMMI e GGE, assumindo uma estrutura fatorial analítica para a matriz de variância-covariância genética. Por se tratar de uma análise de modelo linear misto, eles possuem as vantagens em relação às suas versões os efeitos como fixos, por exemplo, a possibilidade de modelar a heterogeneidade de variância entre os ambientes e a correlação espacial dentro do ambiente, além de lidar com conjuntos de dados incompletos. Além disso, considerar genótipos como aleatórios possibilita a incorporação dos coeficientes de parentesco na forma fatorial analí-

tica oferecendo estimativas mais precisas para valores genéticos (CROSSA et al., 2006; OAKEY et al., 2007; BURGUEÑO et al., 2007; CROSSA et al., 2011).

2.3 O modelo AMMI

O modelo AMMI (*Additive Main Effects and Multiplicative Interaction*) pertence a uma classe mais geral de modelos, denominada modelos lineares-bilineares (*General Linear-Bilinear Model - GLBM*) (CORNELIUS; SEYEDSADR, 1997) que podem ser representados da seguinte forma:

$$Y = X_r\beta_r + X_c\beta_c + \sum_{k=1}^p \lambda_k\alpha_k\gamma_k^T + \epsilon \quad (1)$$

em que: β_r e β_c denotam os vetores de efeitos de linhas e colunas, respectivamente; α_k , γ_k e λ_k os vetores singulares relacionados a linhas e colunas e os valores singulares obtidos por decomposição por valor singular; X_r e X_c são matrizes de delineamentos; e o vetor ϵ contém os erros experimentais (c e r são os números de linhas e colunas referente à matriz de dupla entrada) e ϵ corresponde ao vetor de erros aleatórios.

Além do modelo AMMI, outras variações podem ser obtidas a partir da equação (1). É o caso, por exemplo, do *Rows Regression* (RREG), em que o efeito de linhas é confundido com a interação. Uma variante desse modelo é o GGE-biplot, proposto por Yan et al. (2000), em que apenas os dois primeiros componentes principais são considerados para efeito do estudo de estabilidade e adaptabilidade genotípica. Os modelos AMMI e GGE-biplot são os mais populares dentre os GLBM, com ampla aplicabilidade em programas de melhoramento de plantas e ensaios agrônômicos envolvendo dados multiambientais.

O procedimento AMMI, aqui enfatizado, incorpora componentes aditivos

para os efeitos principais de genótipos e de ambientes e, ainda, componentes multiplicativos para o efeito da interação entre esses fatores, combinando, em uma única abordagem, o estudo da adaptabilidade e estabilidade por meio da análise de componentes principais. Esse é um modelo de efeito fixo proposto por Gollob (1968) e Mandel (1969, 1971), com ampla aplicabilidade em diversas áreas de estudo. Segundo o modelo AMMI, a resposta média de um genótipo i em um ambiente j (Y_{ij}) é obtida por:

$$Y_{ij} = \mu + \tau_i + \delta_j + \sum_{k=1}^p \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \quad (2)$$

em que:

μ é uma constante, geralmente a média geral;

τ_i é o efeito do i -ésimo genótipo, $i = 1, \dots, r$, sendo r o número de linhas (ou genótipos) referentes à tabela com as médias do ensaio;

δ_j é o efeito do j -ésimo ambiente, $j = 1, \dots, c$, sendo c o número de colunas (ou ambientes) referentes à tabela com as médias do ensaio;

λ_k é o k -ésimo valor singular referente à matriz $GE^T GE$ ou $GEGE^T$, com $k = 1, \dots, p$, $\{p \leq \min(r - 1, c - 1)\}$, sendo p o posto da matriz de interação entre genótipos e ambientes (GE);

α_{ik} é o i -ésimo elemento relativo ao genótipo i do k -ésimo vetor singular referente à matriz $GEGE^T$;

γ_{jk} é o j -ésimo elemento relativo ao ambiente j do k -ésimo vetor singular referente à matriz $GE^T GE$;

ε_{ij} é o erro experimental, admitido ser normal e identicamente distribuído com, média zero e variância comum, $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Os termos α_{ik} , γ_{jk} e λ_k são referidos como componentes bilineares ou multiplicativos do modelo e resultam da decomposição por valores singulares (DVS)

da matriz de interação entre genótipos e ambientes, estando sujeitos às restrições para assegurar identificabilidade: $\sum_i \tau_i = \sum_j \delta_j = \sum_i \alpha_{ik} = \sum_j \gamma_{jk} = 0$, e ainda àquelas provenientes da DVS $\sum_i \alpha_{ik} \alpha_{ik'} = \sum_j \gamma_{jk} \gamma_{jk'} = 0$, para $k \neq k'$ e $\sum_i \alpha_{ik} = \sum_j \gamma_{jk}^2 = 1$. Além disso, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ para os valores singulares.

O ajuste do modelo AMMI é realizado em duas etapas. Primeiramente, ajustam-se os efeitos principais pelo método de mínimos quadrados aplicado à matriz de médias $Y_{r \times c} = [Y_{ij}]$. A segunda etapa corresponde à DVS da matriz de interações, que é obtida como resíduo do ajuste aos efeitos principais pela análise de variância. A DVS determina um desdobramento ortogonal para a soma de quadrados da interação ($SQ_{G \times E}$) e a soma acumulativa dos sucessivos termos do desdobramento produz aproximações cada vez melhores para $SQ_{G \times E}$.

Nesse sentido, é possível recuperar uma parte substancial da soma de quadrados a partir de poucos eixos (referida como padrão), mais diretamente associados à interação, e descartar o ruído (considerado a variabilidade) geralmente, inexplicável entre as repetições, o que tem por objetivo melhorar a capacidade de predição do modelo (GAUCH, 1988; ZOBEL; WRIGHT; GAUCH, 1988). Desta forma, o índice k ($k = 1, 2, \dots, p$; $p = \min(r - 1, c - 1)$), referente ao número de termos bilineares, tomado até n no somatório ($n < p$), determina uma aproximação de mínimos quadrados para a matriz de interação (GABRIEL, 1978).

Assim, é possível identificar um termo ρ_{ij} que representa a falta de ajuste do modelo ($\rho_{ij} = \sum_{k=n+1}^p \lambda_k \alpha_{ik} \gamma_{jk}$) e constitui um resíduo adicional na análise, que pode ser descartado a partir de algum critério estatístico, sendo essa uma das maiores virtudes do AMMI ou, mais geral, dos modelos GLBM, em relação aos métodos de análises mais tradicionais.

Outro atrativo dessa abordagem está na possibilidade de plotar, grafica-

mente, os escores genotípicos e ambientais nos eixos principais (Biplots) (GAUCH, 1988; YAN; TINKER, 2006). O princípio fundamental é representar indivíduos e variáveis, simultaneamente, num espaço de baixa dimensão, de forma a facilitar a observação de padrões existentes. Os produtos internos entre vetores compostos pelos escores genotípicos e ambientais dão uma aproximação do efeito da interação e permitem identificar ambientes e genótipos que não contribuem com a interação (estáveis), bem como apontar possíveis combinações favoráveis de genótipos a ambientes específicos (GABRIEL, 1971; GAUCH, 2006; GAUCH; PIEPHO; ANNICCHIARICO, 2008; YAN; KANG, 2002; SILVA; BENIN, 2012).

2.4 Determinação do número de termos bilineares no modelo AMMI

O procedimento de análise AMMI permite o ajuste de um número pequeno de componentes multiplicativos sem, no entanto, perder o poder explicativo, ou seja, permite a obtenção de um modelo parcimonioso. Tal procedimento oferece uma família de modelos em função do número de termos bilineares a serem retidos para explicar a interação, que pode ser determinado por diversos procedimentos.

Os métodos mais utilizados, para esse fim, são baseados em testes F aproximados, dentre os quais se destacam: o teste de Gollob (1968), na qual a significância do k -ésimo valor singular é verificada em relação ao quadrado médio do erro médio, e o teste proposto por Cornelius, Seyedsadr e Crossa (1992) realizado sob hipótese nula de que não há mais do que n termos determinando a interação. Piepho (1995) comparou quatro procedimentos para seleção de termos multiplicativos em modelo AMMI, envolvendo teste F , constatando maior robustez do teste F_R de Cornelius sobre os demais.

Questionamentos sobre a partição do número de graus de liberdades relacionados a cada termo da interação (GAUCH; ZOBEL, 1996), bem como so-

bre a pressuposição de que λ_k^2 seguiria uma distribuição qui-quadrado, coloca incerteza quanto à legitimidade da utilização dos testes F . Nesse sentido, métodos baseados em computação intensiva ganharam aplicabilidade, como aqueles que utilizam procedimentos de validação cruzada (GAUCH, 1988; DIAS; KRZANOWSKI, 2003; GABRIEL, 2002; KRZANOWSKI, 1987).

Outro método para o ajuste do modelo AMMI foi proposto por Cornelius et al. (1996), Cornelius e Crossa (1999) e que se baseia no uso de preditores lineares que produzem encolhimento "*Shrinkage*" nos valores dos componentes que descrevem a interação. Nesse método, não é efetuado qualquer procedimento de seleção do número de eixos principais, e, em geral, produz modelos mais precisos que aqueles obtidos a partir de testes. Entretanto, um problema é que não produz modelos parcimoniosos, no sentido de reduzir a quantidade de parâmetros, além de que, o algoritmo para estimar os valores singulares, frequentemente, viola as restrições do modelo, o que requer a incorporação de passos extras no processo de estimação (SILVA et al., 2015).

2.5 Inferência Bayesiana

A inferência bayesiana está fundamentada segundo uma interpretação, essencialmente, subjetiva de probabilidade em que toda incerteza genérica, ou “grau de crença”, a qualquer quantidade desconhecida pode ser interpretada em um sentido probabilístico, ou seja, toda quantidade desconhecida é encarada como uma variável aleatória (SPIEGELHALTER; ABRAMS; MYLES, 2004). Assim, é possível incorporar e atualizar a informação prévia com as informações fornecidas pelos dados do experimento. Para fazer esta associação, a ferramenta utilizada é o Teorema de Bayes, dado por:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (3)$$

Na expressão (3), $p(y|\theta)$ é a probabilidade conjunta, ou função de verossimilhança, $p(\theta)$ denota a densidade a priori e $p(\theta|y)$ é a distribuição conjunta a posteriori, base para toda a inferência a ser realizada. O termo $p(y)$, por sua vez, representa a densidade de probabilidade associada ao vetor de dados y e é um fator de proporcionalidade que garante que a integral seja igual a 1, sendo expresso por:

$$p(y) = \int_{\theta} p(y|\theta)p(\theta)d\theta.$$

Como os dados são supostamente conhecidos, a distribuição marginal $p(y)$ é uma constante em relação a $p(\theta|y)$ e, usualmente, a regra de Bayes é escrita como uma relação de proporcionalidade (\propto), ou seja, $p(\theta|y) \propto p(y|\theta)p(\theta)$.

2.6 Incorporação de informações a priori

Em estudos envolvendo inferência Bayesiana, a informação que se tem sobre um parâmetro, antes de ele ser acessado, é incorporada à análise por meio de distribuições a priori, que devem refletir o “grau de confiança” do pesquisador sobre os possíveis valores dos parâmetros. A incorporação dessas informações substanciais na análise bayesiana pode ser realizada por diversos critérios. Um deles é utilizar famílias de distribuições conjugadas, que constitui um dos principais conceitos para modelar a distribuição a posteriori pela facilidade de manipulações matemáticas.

Famílias conjugadas baseiam-se na ideia de que as distribuições a posteriori e a priori pertençam à mesma classe de distribuições. Assim, impõe-se que a família de distribuições a priori H seja fechada sob amostragem de $F =$

$\{f(y|\theta) : \theta \in \Theta\}$, em que F denota a família de distribuições que modela a situação abordada. Então, a Família H é conjugada de F , representada por $h(\theta) \in H \Rightarrow h(\theta|y) \propto h(\theta)f(y|\theta) \in H$. Incorporar informação a priori, nesse contexto, se dá pela escolha apropriada dos hiperparâmetros, que correspondem aos parâmetros da priori (BERNARDO; SMITH, 1994).

Caso se tenha um conhecimento vago a respeito do parâmetro, ou não se queira incorporar qualquer informação além daquelas fornecidas pelos dados amostrais (ignorância a priori), é comum utilizar as chamadas prioris não informativas. Dentre elas, destacam-se aquelas obtidas pela regra de Bayes-Laplace, que consideram os valores do parâmetro igualmente prováveis (priori uniforme), e pelo método de Jeffreys (1961), baseado na medida de informação esperada de Fisher (PAULINO; TURKMAN; MURTEIRA, 2003). Existem alternativas para a incorporação de prioris não-informativas. Jaynes (2003), por exemplo, propôs utilizar o critério de máxima entropia na especificação da distribuição a priori e Bernardo (1979) defende, também, a especificação da priori não informativa fundamentada em critérios dentro do contexto da teoria da informação.

Como a densidade da distribuição a posteriori é obtida por meio da informação a priori e dos dados do experimento, ela será equivalente à função de verossimilhança, caso a priori seja não informativa ou o número de observações na amostra seja elevado. Obviamente, tal densidade será substancialmente diferente da função de verossimilhança dependendo da quantidade de informação incorporada ao parâmetro, dada pela priori.

2.7 Estimação dos parâmetros

No contexto bayesiano, a estimação pontual pode ser conduzida com base na teoria da decisão, que envolve a especificação do espaço (Θ) dos possíveis va-

lores para o parâmetro θ , o espaço das ações possíveis (A) e o espaço dos possíveis resultados do experimento Ω (MIGON; GAMERMAN, 1999). Nesse sentido, um bom estimador “ $\delta(X)$ ” é aquele em que o erro $\delta(X) - \theta$ está próximo de zero com alta probabilidade. Então, para cada valor θ é possível estimar $a \in \Theta$ e considera-se a função perda esperada $L(\theta, a)$ a ser considerada, de modo a quantificar a perda obtida no processo de estimação.

Assim, a função $L(\theta, a)$ está associada a cada regra de decisão $\delta(x)$, $x \in \Omega$, para cada possível valor $\theta \in \Theta$. Obtém-se, então, uma medida de penalização relativa à decisão $\delta(x)$ quando o valor verdadeiro do parâmetro é θ . A perda esperada a posteriori é o risco associado a uma determinada regra de decisão $R(\delta)$:

$$R(\delta) = E(\theta|x)(L(\delta, \theta)).$$

A decisão que minimiza o risco esperado, de acordo com a função perda apropriada, é o estimador pontual para o parâmetro. As estimativas pontuais mais frequentemente utilizadas na inferência bayesiana são a média, a mediana e a moda da distribuição conjunta a posteriori, que são obtidas como risco esperado referente às funções de perdas quadráticas, perdas absolutas e zero-um, respectivamente (BERGER, 1985).

Um resumo mais substancial da função conjunta a posteriori é obtido por meio de intervalos de valores no espaço paramétrico, denominados regiões de credibilidade. Uma região $C \subset \Theta$ é um intervalo de credibilidade para θ se $P(\theta \in C|y) \geq 1 - \alpha$, sendo θ uma quantidade desconhecida no espaço paramétrico Θ . O termo $1 - \alpha$ é dito ser o nível de credibilidade. Baseando-se nesta definição, uma infinidade de regiões de credibilidade distintas podem ser construídas, considerando um determinado nível de credibilidade. Em muitas situações

são utilizadas regiões que contenham todos os valores do parâmetro com maior credibilidade a posteriori e são denominadas regiões com máxima densidade a posteriori (*Highest Posterior Density - HPD*) (BOX; TIAO, 2011). Assim, regiões de credibilidade possibilitam uma interpretação probabilística direta em relação ao parâmetro, que é considerado uma variável aleatória (BERNARDO; SMITH, 1994).

Na maioria das aplicações, em inferência bayesiana é necessário integrar funções complexas e multidimensionais. A inferência exata somente será possível se tais integrais puderem ser calculadas analiticamente; caso contrário, métodos de aproximações numéricas serão necessários. Na realidade, este é o principal aspecto envolvido em procedimentos bayesianos. A modelagem de sistemas físicos muitas vezes exige o emprego de modelos hierárquicos sofisticados, em que o tratamento analítico seria muito difícil, ou mesmo impraticável. Todavia, existem métodos numéricos em que as soluções podem ser aproximadas através de procedimentos iterativos, com destaque para os procedimentos de simulações baseados em Cadeias Markov, via método de Monte Carlo (*Markov chain Monte Carlo - MCMC*) (GELMAN et al., 2003).

2.8 Cadeias de Markov e Inferência via técnicas MCMC

Uma cadeia de Markov discreta é um processo estocástico, ou seja, uma sequência de variáveis aleatórias X_n , ($n = 0, 1, 2, \dots$), em que X_n toma valores no conjunto finito $S = 0, 1, \dots, N - 1$ e são chamados de estados da cadeia de Markov. Os índices n podem ser interpretados como estágios ou períodos de tempo. Se $X_n = i$, diz-se que o processo está no estado i no tempo n . Uma cadeia de

Markov deve satisfazer a seguinte propriedade:

$$\Pr(X_n = j | X_{n-1} = i, X_{n-2} = k, \dots, X_0 = m) = \\ \Pr(X_n = j | X_{n-1} = i) = p(i, j);$$

$i, j, k, \dots, m \in S$.

Assim, na cadeia de Markov a distribuição de X_t , dados todos os valores anteriores, X_0, \dots, X_{n-1} , depende apenas do estado imediatamente anterior, ou seja, X_{t-1} .

Como mencionado anteriormente em inferência Bayesiana, procedimentos baseados em Monte Carlo Via Cadeia de Markov são uma das estratégias mais difundidas. Trata-se de um procedimento que possibilita obter amostras dos parâmetros do modelo por meio de um processo iterativo, que por sua vez tenha, pelo menos em um sentido assintótico, a distribuição a posteriori desejada como equilíbrio ou distribuição estacionária. Após um número finito de iterações, espera-se que as cadeias produzidas atinjam a distribuição de equilíbrio e, a partir deste ponto, obtém-se uma amostra da distribuição a posteriori dos parâmetros (GAMERMAN; LOPES, 2006).

Dentre os métodos para se construir cadeias de Markov está o algoritmo de Gibbs, utilizado quando se consegue amostrar diretamente da distribuição condicional completa a posteriori. Já o algoritmo Metropolis-Hastings é utilizado caso as densidades condicionais a posteriori não possuam formas conhecidas ou, mesmo sendo conhecidas, não permitam amostragem direta (PAULINO; TURKMAN; MURTEIRA, 2003).

2.8.1 O algoritmo Metropolis-Hastings

Como já enfatizado, toda inferência bayesiana, é realizada com base na distribuição conjunta a posteriori, que é obtida ao efetuar a conexão das distribuições a priori especificadas para os parâmetros e os dados experimentais, via teorema de Bayes, por manipulações algébricas. Encontrar as densidades marginais é uma tarefa difícil. Contudo, é possível que se conheça o núcleo da distribuição condicional (a menos de uma constante). Nessa nuance, a amostragem pode ser conduzida direta por meio de algum algoritmo.

Entretanto, em muitas situações, a distribuição condicional completa do parâmetro não pode ser identificada, ou ainda, a amostragem não pode ser realizada diretamente. Nessas situações, o algoritmo Metropolis-Hastings pode ser utilizado para gerar sorteios de candidatos de uma distribuição auxiliar. Esses sorteios são, então, “corrigidos” para que, assintoticamente, eles se comportem como observações aleatórias do equilíbrio.

A cadeia de Markov produzida pelo algoritmo Metropolis-Hastings em cada estágio é construída em duas etapas: uma etapa de proposta e uma etapa de aceitação que figuram como os principais aspectos do *kernel* de transição entre os estados da cadeia. Suponha que X_n seja o estado atual da cadeia e que $Y_{n+1} = y$ seja proposto como o próximo estágio, amostrado da distribuição proposta $q(x, \cdot)$. O ponto candidato y é, então, aceito com probabilidade $a(x, y)$. Se o candidato for aceito, o próximo estado da cadeia será o ponto $X_{n+1} = y$; caso contrário, a cadeia não se move, ou seja, $X_{n+1} = x$.

A probabilidade de aceitação é definida como $a(X_n|Y_{n+1}) = \min\left(1, \frac{\pi(Y_{n+1})q(X_n|Y_{n+1})}{\pi(X_n)q(Y_{n+1}|X_n)}\right)$, sendo π a distribuição de interesse.

2.8.2 Amostrador de Gibbs

Dentre os métodos envolvendo técnicas de MCMC, o amostrador de Gibbs vem sendo o mais comumente empregado quando as distribuições condicionais completas permitem amostragem direta. O amostrador de Gibbs fornece um processo iterativo para geração de sucessivas amostras das distribuições condicionais completas. Gamerman e Lopes (2006) apresentam uma descrição desse procedimento.

Seja $\theta^0 = (\theta_1^0, \dots, \theta_d^0)$ um valor arbitrário inicial para o vetor θ . A j -ésima iteração referente à cadeia de Markov é obtida pelo seguinte algoritmo:

1. Inicializar o contador de iteração de cada $j = 1$ e o conjunto de valores iniciais

$$\theta^0 = (\theta_1^0, \dots, \theta_d^0)^\top;$$

2. Obter um novo valor $\theta^j = (\theta_1^j, \dots, \theta_d^j)^\top$ de θ^{j-1} por meio de sucessivas gerações dos valores:

$$\theta_1^j \sim p\left(\theta_1 | \theta_2^{(j-1)}, \dots, \theta_d^{(j-1)}\right),$$

$$\theta_2^j \sim p\left(\theta_2 | \theta_1^{(j)}, \dots, \theta_d^{(j-1)}\right),$$

$$\vdots$$

$$\theta_d^j \sim p\left(\theta_d | \theta_1^{(j)}, \dots, \theta_{d-1}^{(j)}\right).$$

3. Mudar o contador de j para $j + 1$ e repetir o passo 2 até atingir o número de iterações predeterminado.

Quando a convergência é alcançada, o valor resultante θ^j é uma amostra de p . À medida que o número de iterações aumenta, a distribuição de transição se

aproxima do seu estado de equilíbrio e a convergência é assumida de forma aproximada. O amostrador de Gibbs explorando as distribuições condicionais completas por meio de algoritmo iterativo foi proposto inicialmente por Geman e Geman (1984) no contexto de reconstrução de imagens e, posteriormente, utilizado por Gelfand e Smith (1990) para simulação de distribuições a posteriori.

2.9 O modelo AMMI-bayesiano

A aplicação do método bayesiano ao modelo AMMI, sob as restrições inerentes não mostrou ser uma tarefa simples, pois a restrição de ortonormalidade dos vetores singulares impõe grandes dificuldades ao processo de amostragem, já que o suporte da densidade condicional a posteriori não é trivial. Viele e Srinivasan (2000) mostraram como contornar essa dificuldade, conduzindo o processo de amostragem via transformação linear ortogonal utilizando uma variável auxiliar em um “espaço corrigido”. Para tanto, utilizaram técnicas de MCMC através do amostrador de Gibbs com a incorporação de subcadeias obtidas pelo algoritmo Metropolis-Hastings.

Liu (2001) abordou questões práticas e teóricas referentes ao método proposto por Viele e Srinivasan (2000), comparando o AMMI-Bayesiano com versões clássicas do modelo AMMI, incluindo o procedimento proposto por Cornelius e Crossa (1999). Dentre as contribuições do trabalho de Liu (2001), pode-se ressaltar a derivação algébrica das densidades condicionais completas a posteriori para os vetores singulares, que permitiu amostragem direta pelo algoritmo de Gibbs, evitando utilizar passos de Metropolis-Hastings. Isto forneceu maior velocidade e estabilidade ao algoritmo, superando, dessa forma, dificuldades encontradas por Viele e Srinivasan (2000).

Uma versão em notação vetorial do modelo proposto por Viele e Sriniva-

san (2000) é apresentada por Crossa et al. (2011), por meio da seguinte equação:

$$\mathbf{y} = \mathbf{1}_n\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{g} + \sum_{k=1}^t \lambda_k \text{diag}(\mathbf{Z}\boldsymbol{\alpha}_k)\mathbf{X}\boldsymbol{\gamma}_k + \boldsymbol{\varepsilon} \quad (4)$$

Na equação (4), o vetor \mathbf{y} contém $n = bcr$ respostas fenotípicas, sendo r , b e c os números de genótipos, blocos e ambientes, respectivamente. As matrizes \mathbf{X} e \mathbf{Z} são utilizadas para distribuir os efeitos no modelo (matrizes de incidência) e $\mathbf{1}_n$ é o vetor cujas coordenadas têm valor igual a um. Os vetores $\boldsymbol{\beta}$ e \mathbf{g} ($l = bc$) contêm os parâmetros de efeitos principais de blocos e de genótipos, respectivamente. Os termos λ_k , $\boldsymbol{\alpha}_k$ e $\boldsymbol{\gamma}_k$ são os k -ésimos valores e vetores singulares, como no modelo AMMI tradicional, com $k = 1, \dots, p$; sendo $p = \min(r, c)$, o posto da matriz GEI. A média geral é denotada por μ e o vetor $\boldsymbol{\varepsilon}_{n \times 1}$ contém os efeitos dos erros aleatórios, com $\boldsymbol{\varepsilon} \sim N_n(0, \sigma_e^2 \mathbf{I})$.

As densidades a priori para os parâmetros do modelo são as mesmas propostas por Viele e Srinivasan (2000):

$$\mu | \mu_\mu, \sigma_\mu^2 \sim N(\mu_\mu, \sigma_\mu^2)$$

$$\boldsymbol{\beta} | \mu_\beta, \sigma_\beta^2 \sim N(\mu_\beta, \sigma_\beta^2);$$

$$\mathbf{g} | \mu_g, \sigma_g^2 \sim N(\mathbf{0}, \mathbf{I}\sigma_g^2);$$

$$\lambda_k | \mu_{\lambda_k}, \sigma_{\lambda_k}^2 \sim N^+(\mu_{\lambda_k}, \sigma_{\lambda_k}^2);$$

$$\boldsymbol{\alpha}_k \sim \text{uniforme esférica no subespaço corrigido};$$

$$\boldsymbol{\gamma}_k \sim \text{uniforme esférica no subespaço corrigido};$$

$$\sigma_e^2 \sim \text{inv - escalada} - \chi^2(v_e, S_e^2) .$$

Para os parâmetros μ , β , g e λ_k são atribuídas densidades normais. A notação N^+ indica que para λ_k a densidade normal é truncada nos valores positivos e deve satisfazer a condição $\lambda_k > \lambda_{k+1} > 0$. Para os vetores singulares, α_k e γ_k são atribuídas distribuições uniformes esféricas no subespaço corrigido, que são casos particulares da distribuição esférica von Mises-Fisher quando o parâmetro de concentração é igual a zero (MARDIA; KENT; BIBBY, 1979). Para σ_e^2 utiliza-se como densidade a priori uma qui-quadrado escalada invertida.

A distribuição condicional dos dados tem densidade normal multivariada $\mathbf{y}|\mu, \alpha, \gamma, \lambda, g, \beta, \sigma_e^2 \sim N(\boldsymbol{\theta}, \mathbf{I}_n \sigma_e^2)$, em que \mathbf{I}_n é a matriz identidade, $\boldsymbol{\theta} = \mu + \mathbf{X}\beta + \mathbf{Z}g + \sum_{k=1}^t \lambda_k \text{diag}(\mathbf{Z}\alpha_k)\mathbf{X}\gamma_k$ e σ_e^2 é a variância residual.

Nessa abordagem são assumidas todas as pressuposições do modelo AMMI clássico (efeitos fixos). A flexibilidade do modelo AMMI-Bayesiano para incorporar inferência à representação biplot, bem como informações históricas de experimentos foram exemplificadas por Crossa et al. (2011) e Perez-Elizalde, Jarquin e Crossa (2012). Oliveira et al. (2015) mostraram como efeito aleatório para genótipo pode ser obtido por meio da atribuição de uma priori hierárquica, em que se considera incerteza sobre o parâmetro de variância genotípica. Silva et al. (2015), por sua vez, apresentaram uma justificativa Bayesiana para os estimadores Shrinkage propostos por Cornelius e Crossa (1999).

2.10 Critério de seleção de modelos em inferência bayesiana

A seleção de termos bilineares em modelos AMMI-Bayesiano tem sido realizada por meio de critérios de informação como *Akaike Information Criterion* (AIC) e *Bayesian Information Criterion* (BIC), ou também o fator de Bayes (LIU, 2001; PEREZ-ELIZALDE; JARQUIN; CROSSA, 2012). Segundo Liu (2001), critérios de informação fornecem melhores modelos quando se utilizam prioris não

informativas, ao passo que o Fator de Bayes se mostra melhor quando informações são acrescentadas, por meio das prioris, às componentes de variância dos valores singulares. Seguindo a ideia de obtenção de modelos mais parcimoniosos, Silva et al. (2015) propõem utilizar priori informativa para os parâmetros bilineares, no que resulta em um efeito aleatório para valores singulares e conseqüentemente, em um efeito de encolhimento nas estimativas do parâmetro. É esperado que se obtenham modelos de menor dimensionalidade, no sentido de seleção de poucos eixos principais, se algum critério, para esse fim, for utilizado no método proposto por Silva et al. (2015).

Em aplicações práticas, é comum assumir incerteza quanto ao modelo que melhor se ajusta a um determinado conjunto de dados. Dada a importância da escolha adequada do melhor modelo que descreve os dados, vários estudos propondo metodologias (critérios) foram, e ainda são, realizados nesta busca. Dentre essas propostas, pode-se destacar aquelas realizadas por Akaike (1974), Schwarz (1978) e Kass e Raftery (1995).

De forma geral, realizar seleção de modelos significa buscar critérios que satisfaçam as condições de parcimônia e bom ajuste. Dentre as formas mais usuais em inferência bayesiana, destacam-se o Fator de Bayes, o critério de Akaike (AIC) e o critério Bayesiano (BIC). Neste trabalho serão utilizados AIC, BIC e também o Critério de Informação de Akaike Monte Carlo (AICM) como procedimento de seleção de eixos principais a serem mantidos (significativos) no modelo AMMI.

2.10.1 Critério de Informação de Akaike (AIC)

O critério de informação de Akaike, apresentado em Akaike (1974), é uma medida relativa à quantidade de ajuste de um modelo em detrimento ao número de parâmetros que compõe o modelo, admitindo a existência de um único modelo

que descreveria os dados, entre os possíveis em avaliação. É baseado na função de log-verossimilhança penalizada pelo número de parâmetros do modelo e será tido como modelo mais adequado o que obtiver menor valor de AIC entre todos em teste. Sua estatística de teste é dada por:

$$AIC = -2\ln(L(\theta_p)) + 2p$$

sendo $L(\theta_p)$ a máxima verossimilhança do modelo estimado e p o número de parâmetros.

2.10.2 Critério de Informação Bayesiano (BIC)

O Critério de informação Bayesiano (BIC), proposto em Schwarz (1978), é definido em termos de probabilidade a posteriori do modelo ajustado e será considerado como modelo mais adequado o que obtiver menor valor de BIC entre todos em teste. Sua estatística de teste é dada por:

$$BIC = -2\ln(L(\hat{\theta}_p)) + \ln(n)p$$

sendo $L(\hat{\theta}_p)$ a máxima verossimilhança do modelo estimado, p o número de parâmetros e n corresponde ao número de observações contida na cadeia.

2.10.3 Critério de Informação de Akaike Monte Carlo (AICM)

Raftery et al. (2007) propõem o AICM como sendo uma correspondência a posteriori baseada em simulação do critério de seleção do modelo AIC. E esta proposta tem como vantagem de poder ser estimada diretamente a partir da amostra a posteriori obtida pelo método Monte Carlo via Cadeia de Markov. Sendo assim o modelo tido como mais adequado será o que obtiver menor valor de AICM.

E sua estatística de teste pode ser dada por (CHEN; KUO; LEWIS, 2014):

$$AICM = 2 (s_l^2 - \bar{l})$$

sendo: \bar{l} a média a posteriori da log-verossimilhança marginal e s_l^2 a variância a posteriori da log-verossimilhança marginal.

2.11 *Reversible Jump* em cadeias de Markov

Os métodos para se obter o melhor modelo que se ajusta a um conjunto de dados, como os descritos acima, exigem que os modelos candidatos sejam ajustados de forma independente para que, posteriormente, as respectivas respostas sejam confrontadas. Neste contexto, a realização de vários processos MCMC, envolvendo algoritmo de Gibbs ou Metropolis-Hastings dependendo da modelagem em questão, poderia constituir uma tarefa dispendiosa, envolvendo um alto custo computacional, tornando-a até mesmo inviável.

Uma maneira sensata de tratar a seleção de modelos, no contexto bayesiano, seria permitir que o modelo $(k', \theta_{k'})$ dependa dos valores atuais (k, θ_k) , em que k é o indicador do modelo $M_k \{k = 1, 2, \dots, I\}$ e θ_k o vetor paramétrico associado ao modelo (sendo $k' \neq k$). Contudo, para modelos em competição que envolvem vetores com dimensões diferentes, as abordagens convencionais de métodos Monte Carlo via Cadeias de Markov, como o algoritmo Gibbs e o algoritmo Metropolis Hastings, não podem ser usados, já que esses procedimentos de amostragem exigem que o vetor paramétrico tenha dimensão constante.

Uma forma de contornar essa dificuldade foi apresentada em Green (1995), que propôs um método MCMC com *Reversible Jump* (RJMCMC), podendo ser visto como uma generalização do algoritmo Metropolis-Hastings em situações em

que existem espaços paramétricos com dimensionalidade não fixa, (a dimensão de θ_k pode variar entre os estados da cadeia de Markov). Esse procedimento é baseado na criação de uma cadeia de Markov que permite flexibilidade para “saltar” entre modelos com espaços de parâmetros de diferentes dimensões, de forma a garantir que a cadeia seja irredutível e aperiódica (DELLAPORTAS; FORSTER; NTZOUFRAS, 2002).

Será considerado, aqui, o problema geral na seleção entre k modelos em competição para um vetor de dados observados \mathbf{y} . Nesse sentido, existe uma classe de modelos M_k que, de acordo com k , possui uma probabilidade $p(\mathbf{y}|\theta_k) : \theta_k \in \Theta_k$, sendo θ_k e Θ_k , o vetor e o espaço paramétricos, respectivamente, relacionados ao modelo M_k . O espaço paramétrico envolvendo K modelos é dado por $\Theta = \bigcup_k \{k\} \times \Theta_k$ (DELLAPORTAS; FORSTER; NTZOUFRAS, 2002). A implementação do algoritmo de *Reversible Jump* segue o mesmo princípio do método de Metropolis-Hastings, ou seja, propor um novo valor para a cadeia mediante uma probabilidade de aceitação. Em cada iteração ocorre a atualização dos parâmetros, de acordo com o modelo, utilizando MCMC usuais e a atualização da dimensão com *Reversible Jump*.

2.12 A distribuição invariante

Aqui foi adotada uma abordagem, bem como notação, mais simples, baseada em Waagepetersen e Sorensen (2001) e Sorensen e Gianola (2007), sem recorrer a detalhes da teoria da medida como a adotada por Green (1995). Seja π a distribuição de probabilidade conjunta de (K, θ) , em que $K \in \{1, 2, \dots, I\}$ é o indicador do modelo, I é um número inteiro e θ é um vetor real estocástico possivelmente variando de dimensão.

O vetor paramétrico θ assume valores no conjunto Θ , de espaços $\Theta_k =$

R^{n_k} , como definido anteriormente, com $n_k > 1$. Neste sentido, para $K = k$, o vetor θ só pode obter valores em Θ_k , de modo que π é especificado por $p_k = P(K = k)$ e densidade $f(\cdot|K = k)$ em Θ_k . Assim, para $\mathbf{A}_k \subseteq \Theta_k$, a distribuição de probabilidade conjunta de (K, θ) é dada por:

$$\begin{aligned} P(K = k, \theta \in \mathbf{A}_k) &= P(K = k)P(\theta \in \mathbf{A}_k|K = k) \\ &= p_k \int_{\mathbf{A}_k} f(\theta_k|K = k)d\theta_k. \end{aligned}$$

em que a densidade $f(\cdot|K = k)$ é denotada por f_k .

Então, no contexto em que vários modelos estão em competição, p_k representa a probabilidade a posteriori do modelo k (dado $K = k$) e f_k é a densidade a posteriori do vetor θ de dimensão n_k dos parâmetros associados ao modelo k , de forma que na regra de Bayes:

$$p_k f_k \propto \tilde{p}_k h(\theta_k|k) l(\mathbf{y}|k, \theta_k)$$

em que \tilde{p}_k e $h(\theta|k)$ são as probabilidades a priori do modelo K e o vetor paramétrico θ , dado que $K = k$, respectivamente. A expressão $l(\mathbf{y}|k, \theta)$, por sua vez, é a probabilidade conjunta do vetor \mathbf{y} condicionado a $(K = k, \theta = \theta_k)$.

2.13 Gerando a proposta para um novo estado da cadeia de Markov

No contexto do *Reversible Jump*, cada estado X_i da cadeia de Markov é composto por dois componentes $X_i = (K_i, \theta_{K_i})$, em que K_i indica qual modelo está sendo considerado e θ_{K_i} representa o vetor estocástico que assume valores em Θ_{K_i} . Seja (k, θ_k) o valor do estado atual X_n da cadeia de Markov e considera-se um movimento para o valor $(k', \theta_{k'})$ como próximo estado X_{n+1} da cadeia. Nesse sentido, uma proposta $Y_{n+1} = (Y_{n+1}^{ind}, Y_{n+1}^{par})$, em que “ind” refere-se ao modelo

$K = [(k + 1) = k']$, e “par” ao vetor paramétrico θ_{n+1} , é gerada para X_{n+1} . Então, a partir de uma probabilidade definida $p_{k k'}$ em que $\left(\sum_{k'=1}^I p_{k k'} = 1\right)$, a proposta Y_{n+1}^{ind} é aceita ($Y_{n+1}^{ind} = k'$) e $Y_{n+1}^{par} = \theta_{k'}$ é gerada em $\Theta_{k'}$.

A proposta Y_{n+1}^{par} pode ser construída pela aplicação de um mapeamento determinístico $g_{1k k'}$ associando valores prévios θ e um componente aleatório \mathbf{U} , isto é: $Y_{n+1}^{par} = g_{1k k'}(\theta_k, \mathbf{U})$. Nessa expressão, \mathbf{U} é um vetor aleatório com densidade $g_{k k'}(\theta_k, \cdot)$ em $R^{n_{k k'}}$. A proposta Y_{n+1} é finalmente aceita com uma probabilidade de aceitação $a_{k k'}(\theta_k, Y_{n+1}^{par})$. Os detalhes sobre a probabilidade de aceitação serão abordados mais adiante no texto.

Assim, propondo um movimento a partir do estado atual da cadeia (k, θ_k) para o próximo estado $(k', \theta_{k'}) = (k', g_{1k k'}(\theta_k, u))$, e ainda considerando um movimento na direção oposta de $(k', \theta_{k'})$ para $(k, \theta_k) = (k, g_{1k' k}(\theta_{k'}, u'))$, os vetores (k, θ_k) e $(k', \theta_{k'})$ devem ser de mesma dimensão, ou seja, cumprir a denominada condição de “correspondência de dimensão” ($n_k + n_{k k'} = n_{k'} + n_{k' k}$). Além disso, assume-se que existem funções $g_{2k k'}$ e $g_{2k' k}$ de forma que o mapeamento $g_{k k'}$ dado por $(\theta_{k'}, u') = g_{k k'}(\theta_k, u) = (g_{1k k'}(\theta_k, u), g_{2k k'}(\theta_k, u))$, seja “um a um” (uma bijeção) e ainda que $g_{k k'}$ seja diferenciável, sendo:

$$(\theta_k, u) = g_{k k'}^{-1}(\theta_{k'}, u') = g_{k' k}(\theta_{k'}, u') = (g_{1k' k}(\theta_{k'}, u'), g_{2k' k}(\theta_{k'}, u')).$$

2.13.1 Condição de reversibilidade

Suponhamos uma cadeia de Markov ergódica com espaço de estado S , que converge para uma distribuição invariante π . Considere que $x \in S$ denota o estado atual do sistema e $y \in S$ denota o estado no próximo passo. Seja $p(x, y)$ a probabilidade de uma transição de x para y e que $p(y, x)$ seja a probabilidade de uma transição na direção oposta. Uma cadeia de Markov é considerada reversível

se satisfaz a seguinte condição: $\pi(x)p(x, y) = \pi(y)p(y, x)$; $x, y \in S$.

A condição de reversibilidade temporal para o *kernel* de transição de uma cadeia de Markov, para assegurar a convergência para a distribuição estacionária desejada, pode ser estendida para espaços de estado mais gerais, ou seja, para as cadeias de Markov transdimensionais (SISSON, 2005; ROBERTS; ROSENTHAL, 2006; TIERNEY, 1994). Na proposta transdimensional, como considerada aqui, assumindo que $X_n = (K_n, \theta_n) \sim \pi$, a condição de reversibilidade é dada por:

$$P(K_n = k, \theta_n \in A_k, K_{n+1} = k', \theta_{n+1} \in B_{k'}) = P(K_n = k', \theta_n \in B_{k'}, K_{n+1} = k, \theta_{n+1} \in A_k) \quad (5)$$

para todo $k, k' \in \{1, 2, \dots, I\}$ e todos os subconjuntos A_k e $B_{k'}$ de Θ_k e $\Theta_{k'}$, respectivamente. Na equação (5), o primeiro membro pode ser escrito da seguinte forma:

$$P(K_n = k, \theta_n \in A_k, K_{n+1} = k', \theta_{n+1} \in B_{k'}) = \int_{A_k} P(K_{n+1} = k', \theta_{n+1} \in B_{k'} | X_n = (k, \theta_k)) p_k f_k(\theta_k) d\theta_k, \quad (6)$$

em que o termo $P(K_{n+1} = k', \theta_{n+1} \in B_{k'} | X_n = (k, \theta_k))$ representa o *kernel* de transição. Além disso, pode-se denotar por $Q_{kk'}^a(\theta_k, B_{k'})$ a probabilidade conjunta de aceitação da proposta $Y_{n+1} = (Y_{n+1}^{ind} = k', Y_{n+1}^{par} \in B_{k'})$, condicionada no fato de que o estado atual da cadeia de Markov seja $X_n = (k, \theta_k)$, sendo expressa por:

$$Q_{kk'}^a(\theta_k, B_{k'}) = P(Y_{n+1}^{ind} = k', Y_{n+1}^{par} \in B_{k'} \text{ e } Y_{n+1} \text{ é aceito} | X_n = (k, \theta_k)).$$

Nesse contexto, a probabilidade de rejeitar a proposta é denotada por $s_k(\theta_k) = P(Y_{n+1} \text{ é rejeitado} | X_n = (k, \theta_k))$. Assim, o *kernel* de transição é dado por:

$$P(K_{n+1} = k', \theta_{n+1} \in B_{k'} | X_n = (k, \theta_n)) = Q_{kk'}^a(\theta_k, B_{k'}) + s_k(\theta_k) \times \\ \times I(k = k', \theta_k \in B_{k'})$$

Substituindo no primeiro membro de (5) a expressão dada em (6), em termos do *kernel* de transição obtido anteriormente, tem-se:

$$p_k \int_{A_k} Q_{kk'}^a(\theta_k, B_{k'}) f_k(\theta_k) d\theta_k + p_k \int_{A_k} s_k(\theta_k) I(k = k', \theta_k \in B_{k'}) f_k(\theta_k) d\theta_k \\ = p_k \int_{A_k} Q_{kk'}^a(\theta_k, B_{k'}) f_k(\theta_k) d\theta_k + \\ + p_k \int s_k(\theta_k) I(k = k', \theta_k \in A_k \cap B_{k'}) f_k(\theta_k) d\theta_k \quad (7)$$

sendo:

$$p_k \int_{A_k} Q_{kk'}^a(\theta_k, B_{k'}) f_k(\theta_k) d\theta_k = \\ P(K_n = k, \theta_n \in A_k, Y_{n+1}^{ind} = k', Y_{n+1}^{par} \in B_{k'}, Y_{n+1} \text{ é aceito}).$$

De maneira análoga (por simetria), o segundo membro de (5) pode ser escrito como:

$$\begin{aligned}
p_{k'} \int_{B_{k'}} Q_{k'k}^a(\theta_{k'}, A_k) f_{k'}(\theta_{k'}) d\theta_{k'} + \\
+ p_{k'} \int s_{k'}(\theta_{k'}) I(k = k', \theta_{k'} \in B_{k'} \cap A_k) f_{k'}(\theta_{k'}) d\theta_{k'}.
\end{aligned} \tag{8}$$

Os segundos termos em (7) e (8) são iguais quando $k \neq k'$ (caso em que eles são zero, pela função indicadora), e quando $k = k'$ (nesse caso, o movimento está dentro do mesmo modelo e ambas as expressões são idênticas). Portanto, uma condição suficiente para manter a reversibilidade, para todos os k e k' , é que a seguinte igualdade seja satisfeita:

$$p_k \int_{A_k} Q_{kk'}^a(\theta_k, B_{k'}) f_k(\theta_k) d\theta_k = p_{k'} \int_{B_{k'}} Q_{k'k}^a(\theta_{k'}, A_k) f_{k'}(\theta_{k'}) d\theta_{k'}. \tag{9}$$

para todos conjuntos $A \times B \subset \Theta$ para um espaço de estado geral Θ (GREEN; MIRA, 2001; TIERNEY, 1998; SISSON, 2005).

2.13.2 Probabilidade de aceitação

De uma forma mais explícita, considerando agora a equação (9), em que Y_{n+1}^{ind} é definido como k' com probabilidade $p_{kk'}$. A proposta Y_{n+1}^{par} é, então, gerada em $\Theta_{k'}$, e $Y_{n+1}^{par} \in B_{k'}$ se, e somente se, $g_{1kk'}(\theta_k, \mathbf{U}) = \theta_{k'} \in B_{k'}$, tal como já apresentado na seção anterior. Além disso, a proposta Y_{n+1} é aceita com probabilidade $a_{kk'}(\theta_k, g_{1kk'}(\theta_k, \mathbf{U})) = a_{kk'}(\theta_k, \theta_{k'})$ e $\mathbf{U} \sim q_{kk'}(\theta_k, \cdot)$. Assim, tem-se que:

$$Q_{kk'}^a(\theta_k, B_{k'}) = p_{kk'} \int I(\theta_{k'} \in B_{k'}) a_{kk'}(\theta_k, \theta_{k'}) q_{kk'}(\theta_k, u) du. \tag{10}$$

Então, o lado esquerdo de (9) pode ser desenvolvido da seguinte forma:

$$\begin{aligned}
& p_k \int_{A_k} Q_{kk'}^a(\theta_k, B_{k'}) f_k(\theta_k) d\theta_k \\
&= p_k \int_{A_k} \int I(\theta_{k'} \in B_{k'}) p_{kk'} a_{kk'}(\theta_k, \theta_{k'}) q_{kk'}(\theta_k, u) f_k(\theta_k) d\theta_k du \\
&= p_k \iint I(\theta_k \in A_k, \theta_{k'} \in B_{k'}) p_{kk'} a_{kk'}(\theta_k, \theta_{k'}) q_{kk'}(\theta_k, u) f_k(\theta_k) d\theta_k du \quad (11)
\end{aligned}$$

A partir de um desenvolvimento análogo, por simetria, o lado direito (9) também pode ser reescrito como:

$$\begin{aligned}
& p_{k'} \int_{B_{k'}} Q_{k'k}^a(\theta_{k'}, A_k) f_{k'}(\theta_{k'}) d\theta_{k'} \\
&= p_{k'} \iint I(\theta_{k'} \in B_{k'}, \theta_k \in A_k) p_{k'k} a_{k'k}(\theta_{k'}, \theta_k) q_{k'k}(\theta_{k'}, u') f_{k'}(\theta_{k'}) d\theta_{k'} du'. \quad (12)
\end{aligned}$$

Visando estudar as condições que satisfazem a reversibilidade e, portanto, a igualdade entre (11) e (12), ambas as equações serão expressas como funções das mesmas variáveis. Isso é possível devido às suposições feitas inicialmente, de que o mapeamento é uma bijeção e as funções são diferenciáveis, bem como a condição de correspondência de dimensões.

Usando o fato de que $d\theta_{k'} du' = |\det(g'_{kk'}(\theta_k, u))| d\theta_k du$, sendo $g'_{kk'}(\theta_k, u)$ a derivada parcial de $g_{kk'}$ com relação às variáveis θ_k e u , como representada abaixo:

$$g'_{kk'}(\theta_k, u) = \frac{\partial g_{kk'}(\theta_k, u)}{\partial(\theta_k, u)} = \begin{bmatrix} \frac{\partial g_{1kk'}(\theta_k, u)}{\partial \theta_k} & \frac{\partial g_{2kk'}(\theta_k, u)}{\partial \theta_k} \\ \frac{\partial g_{1kk'}(\theta_k, u)}{\partial u} & \frac{\partial g_{2kk'}(\theta_k, u)}{\partial u} \end{bmatrix}$$

a equação expressa por (12) pode ser reescrita da seguinte maneira:

$$\begin{aligned}
& P(K_n = k', \theta_n \in B_{k'}, K_{n+1} = k, \theta_{n+1} \in A_k) \\
&= \iint I(\theta_{k'} \in B_{k'}, \theta_k \in A_k) p_{k'k} a_{k'k}(\theta_{k'}, \theta_k) \times \\
&\quad \times q_{k'k}(\theta_{k'}, u') p_{k'k} f_{k'}(\theta_{k'}) |det(g'_{kk'}(\theta_k, u))| d\theta_k du.
\end{aligned} \tag{13}$$

Por inspeção, é possível observar que a igualdade entre (12) e (13) é satisfeita se:

$$\begin{aligned}
p_{kk'} a_{kk'}(\theta_k, \theta_{k'}) q_{kk'}(\theta_k, u) p_k f_k(\theta_k) = \\
p_{k'k} a_{k'k}(\theta_{k'}, \theta_k) q_{k'k}(\theta_{k'}, u') p_{k'k} f_{k'}(\theta_{k'}) |det(g'_{kk'}(\theta_k, u))|.
\end{aligned}$$

Escolhendo a probabilidade de aceitação tão grande quanto possível, sujeita à condição detalhada de equilíbrio, chega-se a conclusão de que:

$$a_{kk'}(\theta_k, \theta_{k'}) = \min \left(1, \frac{p_{k'k} q_{k'k}(\theta_{k'}, u') p_{k'k} f_{k'}(\theta_{k'})}{p_{kk'} q_{kk'}(\theta_k, u) p_k f_k(\theta_k)} \left| det \left(\frac{\partial g_{kk'}(\theta_k, u)}{\partial(\theta_k, u)} \right) \right| \right) \tag{14}$$

sendo que $p_{kk'} q_{kk'}(\theta_k, u) p_k f_k(\theta_k) > 0$ e $(\theta_{k'}, u') = g_{kk'}(\theta_k, u)$.

Na prática, $p_{kk'} q_{kk'}(\theta_k, u) p_k f_k(\theta_k) = 0$ só acontece se a cadeia de Markov é inicializada em um estado (k, θ_k) para o qual $p_k f_k(\theta_k) = 0$. Para obter a probabilidade de aceitação para um movimento contrário, ou seja, de $\theta_{k'}$ para θ_k basta calcular o inverso de (14).

2.13.3 Proposta determinística

Como já enfatizado no decorrer do presente texto, no contexto de cadeias de Markov transdimensionais seria sensato permitir que os valores de $(k', \theta_{k'})$ dependam dos valores atuais da cadeia (k, θ_k) da forma mais flexível possível. Nesse sentido, pode ser mais simples aplicar propostas determinísticas para uma mudança de Θ_k para $\Theta_{k'}$, de forma que $Y_{n+1} = g_{1kk'}(\theta_k)$, e ainda usar uma proposta estocástica para o movimento na direção oposta. Nesse caso, a condição de correspondência de dimensão seria igual a $n_k = n_{k'} + n_{k'k}$, desde que $n_{kk'} = 0$. As equações para o mapeamento podem ser reescritas como:

$$\begin{cases} (\theta_{k'}, u') = g_{kk'}(\theta_k) = (g_{1kk'}(\theta_k), g_{2kk'}(\theta_k)) \\ (\theta_k) = g_{kk'}^{-1}(\theta_{k'}, u') = g_{k'k}(\theta_{k'}, u') = g_{1k'k}(\theta_{k'}, u). \end{cases}$$

Observa-se, assim, que a mudança do estado θ_k para o estado $\theta_{k'}$, não envolve a geração de um vetor estocástico \mathbf{U} , ou seja, o movimento é completamente determinístico a partir de um mapeamento. O movimento na direção oposta, por outro lado, requer a geração de um vetor estocástico \mathbf{U}' , nesses casos diz-se que o movimento é estocástico. A condição de reversibilidade, agora é dada por:

$$\begin{aligned} Q_{kk'}^a(\theta_k, B_{k'}) &= p_{kk'} I(g_{1kk'}(\theta_k) \in B_{k'}) a_{kk'}(\theta_k, g_{1kk'}(\theta_k)) \\ &= p_{kk'} I(\theta_{k'} \in B_{k'}) a_{kk'}(\theta_k, \theta_{k'}). \end{aligned} \quad (15)$$

Substituindo no primeiro membro de (9) a expressão $Q_{kk'}^a(\theta_k, B_{k'})$ pelo segundo membro obtido da equação (15), a partir de desenvolvimentos, semelhan-

tes aos apresentados na seção anterior, obtém-se:

$$\begin{aligned}
 & p_k \int_{A_k} Q_{kk'}^a(\theta_k, B_{k'}) f_k(\theta_k) d(\theta_k) \\
 &= \int I(\theta_k \in A_k, \theta_{k'} \in B_{k'}) p_{kk'} a_{kk'}(\theta_k, \theta_{k'}) p_k f_k(\theta_k) d\theta_k. \quad (16)
 \end{aligned}$$

Com uma proposta estocástica para o movimento oposto, o lado direito de (9) fica inalterado e é dado por (12), e $d\theta_{k'} du'$ agora é dado por:

$$d\theta_{k'} du' = |\det(g'_{kk'}(\theta_k))| d\theta_k, \quad (17)$$

sendo

$$g'_{kk'}(\theta_k) = \frac{\partial g_{kk'}(\theta_k)}{\partial \theta_k}.$$

Então, substituindo (17) em (12), e usando (16) obtém-se a expressão para a equação detalhada do equilíbrio:

$$\begin{aligned}
 & \int I(\theta_k \in A_k, \theta_{k'} \in B_{k'}) p_{kk'} a_{kk'}(\theta_k, \theta_{k'}) p_k f_k(\theta_k) d\theta_k \\
 &= \int I(\theta'_k \in B_{k'}, \theta_k \in A_k) p_{k'k} a_{k'k}(\theta'_k, \theta_k) q_{k'k}(\theta'_k, u') \times \\
 & \quad \times p_{k'} f_{k'}(\theta_{k'}) |\det(g'_{kk'}(\theta_k))| d\theta_k \quad (18)
 \end{aligned}$$

sendo $u' = g_{2kk'}(\theta_k)$ uma função de θ_k . Dessa forma, a probabilidade de aceitação é dada agora por:

$$a(\theta_k, \theta_{k'}) = \min \left(1, \frac{p_{k'k} q_{k'k}(\theta'_k, u') p_{k'} f_{k'}(\theta'_{k'})}{p_{kk'} p_k f_k(\theta_k)} |\det(g'_{kk'}(\theta_k))| \right). \quad (19)$$

2.13.4 Mapeamento Identidade

No desenvolvimento apresentado até aqui, a mudança de (k, θ_k) para $(k', \theta_{k'})$ foi proposta por meio de um mapeamento determinístico, como explicitado na seção (2.13.3). Em diversas situações os valores $\theta_{k'}$ podem ser determinados de forma direta, quando todos os parâmetros do modelo proposto são gerados a partir da distribuição proposta, ou seja, a amostra é independente, então $(\theta_{k'}, u) = (\theta_k, u)$. Assim, o jacobiano da transformação não aparece como na expressão (19). Neste sentido, considera-se $g_{1k' k'}(\theta_k, \mathbf{U})$ como um mapeamento identidade, ou seja, $U = \theta_{k'}$, o que resulta em $Y_{n+1}^{par} = \theta_{k'}$.

Pode-se, então, definir simplesmente $\theta_{k'} = u$ e $u' = \theta_k$ no mapeamento que é denotado por $(\theta_{k'}, u) = g_{k' k'}(\theta_k, u) \Rightarrow g_{k' k'}^{-1}(\theta_{k'}, u) = g_{k' k'}(\theta_{k'}, u)$. Assim, $g_{k' k'}(\cdot, \cdot)$ define um mapeamento identidade (LOPES, 2006; LUNN; BEST; WHITTAKER, 2008):

$$\begin{bmatrix} \theta_{k'} \\ u \end{bmatrix} = \mathbf{I}_{(\dim_k + \dim_k)} \times \begin{bmatrix} u \\ \theta_k \end{bmatrix}. \quad (20)$$

Nesse sentido, o vetor aleatório $\theta_{k'}$ é gerado a partir da densidade $q_{k' k'}(\theta_k, \cdot)$ em $\mathbb{R}_{k'}^n$, que pode depender do valor atual θ_k .

Um caminho natural para permitir que o vetor paramétrico $\theta_{k'}$, de um modelo proposto, dependa do vetor paramétrico θ_k , do modelo atual, seria sugerir a distribuição da proposta em espaço amostral de menor dimensão que o de $\theta_{k'}$. Isso pode ser feito no contexto em que $\theta_{k'}$ é aninhado em θ_k ; então podemos propor $\theta_{k'}$ como uma função determinística de θ_k . A probabilidade de gerar e aceitar a proposta $Y_{n+1}^{par} = \theta_{k'}$ é agora dada por:

$$Q_{kk'}^a = p_{kk'} \int I(k' \in B_{k'}) a_{kk'}(\theta_k, \theta_{k'}) q_{kk'}(\theta_k, \theta_{k'}) d\theta_{k'}.$$

Sob essa estratégia, a probabilidade de aceitação é:

$$a_{kk'}(\theta_k, \theta_{k'}) = \min \left(1, \frac{p_{k'k} q_{k'k}(\theta_{k'}, \theta_k) p_{k'} f_{k'}(\theta_{k'})}{p_{kk'} q_{kk'}(\theta_k, \theta_{k'}) p_k f_k(\theta_k)} \right)$$

que, como pode ser visto, não inclui o termo jacobiano decorrente do uso do mapeamento identidade, pois $|\mathbf{J}| = 1$ (LUNN; BEST; WHITTAKER, 2008).

$$g'_{kk'}(\theta_k, u) = \frac{\partial g_{kk'}(\theta_k, u)}{\partial(\theta_k, u)} = \det \begin{bmatrix} \frac{\partial g_{1kk'}(\theta_k, u)}{\partial \theta_k} & \frac{\partial g_{2kk'}(\theta_k, u)}{\partial \theta_k} \\ \frac{\partial g_{1kk'}(\theta_k, u)}{\partial \theta_{k'}} & \frac{\partial g_{2kk'}(\theta_k, u)}{\partial \theta_{k'}} \end{bmatrix}. \quad (21)$$

Considerando $g_{1kk'}$ e $g_{2kk'}$ como em (20), e aplicando as derivadas parciais em (21) obtém-se:

$$\frac{\partial g_{kk'}(\theta_k, u)}{\partial(\theta_k, u)} = \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1 \Rightarrow \left| \frac{\partial g_{kk'}(\theta_k, u)}{\partial(\theta_k, u)} \right| = 1.$$

Além disso, nos casos em que as densidades a posteriori $p(\theta_k|y, k)$ estejam disponíveis em forma fechada, para cada modelo, as mesmas podem ser utilizadas como propostas, ou seja, $q(u'|\theta_{k'}, k', k) = p(\theta_{k'}|y, k)$ e a probabilidade de aceitação, usando o fato de que $p(y|\theta_k, k)p(\theta_k)p(k) = p(\theta_k, k|y)p(y|k)$, se reduz a:

$$a_{kk'}(\theta_k, \theta_{k'}) = \min \left(1, \frac{p(k')p(y|k')p_{k'k}}{p(k)p(y|k)p_{kk'}} \right).$$

Essa expressão se baseia no fato de que a densidade preditiva ou constante de normalização, $p(y|k)$, também é conhecida. Além disso, se $p_{k'k} = p_{kk'}$ a

probabilidade de aceitação é o mínimo entre um e a razão entre a posteriori do modelo indicado por k' e k , isto é, o movimento é automaticamente aceito quando o modelo k' tem maior probabilidade a posteriori em relação ao modelo indicado por k (LOPES, 2006).

3 MATERIAL E MÉTODOS

3.1 Material

Foi simulado um conjunto de dados com 20 genótipos (G1-G20), avaliados em nove ambientes (E1-E9), usando o delineamento de blocos ao acaso, com três repetições. Os genótipos foram separados em três grupos formados por: cinco genótipos com efeito marginal positivo para interação para todos os ambientes, o segundo grupo com cinco genótipos com efeito marginal negativo para interação (genótipos instáveis) e o último grupo formado por 10 genótipos com efeito marginal nulo para interação (genótipos estáveis). O primeiro e segundo grupos de genótipos (genótipos instáveis) foram ambos simulados a partir de uma distribuição normal com média zero e variância 16 considerando o módulo do efeito para as duas situações. Para o último grupo de genótipos foi utilizada uma distribuição normal com média zero e variância igual a um. A simulação foi realizada no software R.

3.2 Métodos

O modelo AMMI, em notação vetorial, pode ser descrito como:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{Z}\mathbf{g} + \sum_{k=1}^t \lambda_k \text{diag}(\mathbf{Z}\boldsymbol{\alpha}_k)\mathbf{X}_2\boldsymbol{\gamma}_k + \boldsymbol{\epsilon} \quad (22)$$

sendo \mathbf{y} o vetor de indivíduos, contendo n observações; $\boldsymbol{\beta}$ e \mathbf{g} são os vetores de efeitos principais de ambiente confundido com as repetições e genotípico, respectivamente; λ_k , $\boldsymbol{\alpha}_k$ e $\boldsymbol{\gamma}_k$ correspondem ao k -ésimo valor singular e os k -ésimos vetores singulares relacionados a ambiente e genótipo, respectivamente, com $k = 1, \dots, t$, e sendo t o posto da matriz GEI. As matrizes \mathbf{X}_1 , \mathbf{X}_2 e \mathbf{Z} correspondem

às matrizes de delineamento associadas aos efeitos β , α_k e \mathbf{g} , respectivamente; e ϵ é o vetor erro aleatório com $\epsilon \sim \mathcal{N}(0, \sigma_e^2 \mathbf{I})$.

No modelo (22) o termo k especifica o número de eixos que são incluídos no modelo que seria de $k = 1$, em que apenas um eixo está sendo incorporado ao modelo até $k = t$, que consiste no modelo completo. Contudo, na abordagem aqui presente o espaço dos modelos é $t + 1$, de forma que poderia se pensar em um índice $k = 0$, em que o modelo correspondente seria $\mathbf{y} = \mathbf{X}_1 \beta + \mathbf{Z} \mathbf{g} + \epsilon$, ou seja, o modelo desconsiderando a parte bilinear.

A função de verossimilhança é dada pela seguinte expressão:

$$L(\bar{\theta}|\mathbf{y}) = (2\pi)^{-\frac{n}{2}} |\mathbf{I}\sigma_e^2|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{y} - \theta)^\top (\mathbf{y} - \theta) \right\} \quad (23)$$

sendo $\theta = \mathbf{X}\beta + \mathbf{Z}\mathbf{g} + \sum_{k=1}^t \lambda_k \text{diag}(\mathbf{Z}\alpha_k)\mathbf{X}\gamma_k$ e $\bar{\theta} = (\beta, \mathbf{g}, \lambda_k, \alpha_k, \gamma_k, \sigma_e^2)$.

3.3 Especificação das distribuições a priori para os parâmetros do modelo

As distribuições a priori para os parâmetros do modelo serão as seguintes:

$$\beta | \mu_\beta, \sigma_\beta^2 \sim N(\mu_\beta, \sigma_\beta^2) \quad \sigma_\beta^2 \rightarrow \infty \quad \beta \sim \text{constante}$$

$$\mathbf{g} | \mu_g, \sigma_g^2 \sim N(0, \mathbf{I}\sigma_g^2)$$

$$\lambda_k | \mu_{\lambda_k}, \sigma_{\lambda_k}^2 \sim N^+(0, \sigma_{\lambda_k}^2)$$

$$\alpha_k \sim \text{Distribuição uniforme esférica no subespaço correto } (SUN I_d)$$

$$\gamma_k \sim \text{Distribuição uniforme esférica no subespaço correto } (SUN I_d)$$

$$\sigma_g^2 \sim \text{Inv-escalada} - \chi^2(v_g, S_g^2) \longrightarrow \frac{1}{\sigma_g^2}$$

$$\sigma_e^2 \sim \text{Inv-escalada} - \chi^2(v_e, S_e^2) \longrightarrow \frac{1}{\sigma_e^2}$$

$$t \sim \text{Poisson-truncada}(\mu)$$

$$\mu \sim \text{gama}(\tau, \nu)$$

Para o parâmetro de escala do valor singular foram estabelecidas três hipóteses acerca da especificação da priori. São elas:

Hipótese a: assume-se que as variâncias dos valores singulares possuem valores grandes (1×10^8), ou seja, não é considerada incerteza com relação a $\sigma_{\lambda_k}^2$ o que corresponde a assumir uma priori constante para λ_k , que é a hipótese apresentada em Oliveira et al. (2015), *Bayesian AMMI* (BAMMI):

Hipótese b: assume-se para $\sigma_{\lambda_k}^2$ uma priori qui-quadrado escalada invertida com parâmetro de escala igual a zero, e o grau de liberdade igual $\Delta = -\frac{(n_{\lambda_k}-1)}{2}$ com a restrição de $0 < \Delta < 1/2$, resultando em uma posteriori com um efeito de encolhimento para os valores singulares, como apresentado em Silva et al. (2015), *Bayesian AMMI Shrinkage* (BAMMIS):

$$\sigma_{\lambda_k}^2 \sim \text{Esc} - \chi^{-2}(v_{\lambda_k}, S_{\lambda_k}^2); \quad \sigma_{\lambda_k}^2 \sim (\sigma_{\lambda_k}^2)^{(\Delta-1)}$$

Hipótese c: assume-se para $\sigma_{\lambda_k}^2$ uma priori gama invertida com grau de liberdade igual a um ($a = 1$) e parâmetro de escala igual a zero ($b = 0$), sendo essa escolha baseada no conceito de Máxima Entropia (ME). Obtém-se assim, uma justificativa para o efeito de encolhimento, sem que seja necessária a imposição de restrições na priori, como apresentado em Silva et al. (2015), *Bayesian AMMI Entropy* (BAMMIE):

$$\sigma_{\lambda_k}^2 \sim Inv - Gama(a, b).$$

Como especificado, para o vetor β foi atribuída uma distribuição normal multivariada com média zero e variância infinita, o que corresponde a uma priori não informativa (denotada por uma constante). Para o vetor \mathbf{g} atribui-se uma distribuição normal multivariada com média zero e variância σ_g^2 . Dada a incerteza acerca de σ_g^2 , a priori referente a \mathbf{g} pode ser considerada informativa. Para as componentes de variâncias σ_e^2 e σ_g^2 foram atribuídas distribuições a priori qui-quadrado escaladas invertidas com graus de liberdade e parâmetros de escala iguais a zero, o que corresponde à priori de Jeffrey para ambos os parâmetros, $\frac{1}{\sigma_e^2}$ e $\frac{1}{\sigma_g^2}$.

As distribuições a priori dos parâmetros bilineares devem satisfazer as restrições do modelo. Desta forma, para os vetores singulares são atribuídas distribuições uniformes esféricas no subespaço corrigido, devido à restrição de ortonormalidade em relação aos vetores (a justificativa de tal atribuição é elucidada em Viele e Srinivasan (2000)). Para os valores singulares λ_k são atribuídas prioris normais truncadas, de modo a satisfazer a restrição $\lambda_k \geq \lambda_{k+1} \geq 0$, com média zero, e para variância λ_k que, como referido, foi submetida a três hipóteses de priori.

Além disso, o número de eixos principais (t) presente no modelo é uma variável aleatória cuja incerteza a priori foi declarada por meio da atribuição de uma distribuição de Poisson truncada com média μ . Para o parâmetro μ foi atribuída uma distribuição Gama com parâmetro de escala e grau de liberdade iguais a um.

3.4 Distribuições condicionais completas a posteriori para os parâmetros do modelo

Combinando a função de verossimilhança com as distribuições a priori estabelecidas, utilizando o teorema de Bayes, obtém-se a distribuição conjunta a posteriori:

$$P(\Phi, |\mathbf{y}) \propto L(\phi, \theta | \mathbf{y}) P(\boldsymbol{\beta}) P(\mathbf{g}) P(\sigma_g^2) P(\sigma_e^2) \times \quad (24)$$

$$\times \left\{ \prod_{k=1}^t P(\lambda_k) P(\boldsymbol{\alpha}_k) P(\gamma_k) \right\} P(t | \mu) P(\mu)$$

em que: $\Phi = (\boldsymbol{\beta}, \mathbf{g}, \sigma_g^2, \sigma_e^2, \lambda_k, \sigma_{\lambda_k}^2, \boldsymbol{\alpha}_k, \gamma_k, t, \mu)$.

As distribuições condicionais completas a posteriori são obtidas a partir da distribuição conjunta a posteriori, que contém todas as informações disponíveis para eventuais inferências sobre a população.

Distribuição condicional completa a posteriori para $\boldsymbol{\beta}$

A distribuição condicional completa a posteriori para $\boldsymbol{\beta}$ é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a $\boldsymbol{\beta}$:

$$P(\boldsymbol{\beta} | \dots) \propto |\mathbf{I}\sigma_e^2|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta}) \right\} P(\boldsymbol{\beta})$$

$$P(\boldsymbol{\beta} | \dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{A}_1 - \mathbf{X}_1\boldsymbol{\beta})^\top (\mathbf{A}_1 - \mathbf{X}_1\boldsymbol{\beta}) \right\} \quad (25)$$

Desenvolvendo a forma quadrática no numerador da expressão (25) entre as chaves, obtém-se:

$$(\mathbf{A}_1 - \mathbf{X}_1\boldsymbol{\beta})^\top (\mathbf{A}_1 - \mathbf{X}_1\boldsymbol{\beta}) = \mathbf{A}_1^\top \mathbf{A}_1 - \mathbf{A}_1^\top \mathbf{X}_1\boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathbf{X}_1^\top \mathbf{A}_1 + \boldsymbol{\beta}^\top \mathbf{X}_1^\top \mathbf{X}_1\boldsymbol{\beta}.$$

A partir de cálculos algébricos e, ainda, observando que termos não dependentes de $\boldsymbol{\beta}$ podem ser absorvidos pela constante de normalização, chega-se a expressão:

$$P(\boldsymbol{\beta}|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left(-2\boldsymbol{\beta}^\top \mathbf{X}_1^\top \mathbf{A}_1 + \boldsymbol{\beta}^\top \mathbf{X}_1^\top \mathbf{X}_1\boldsymbol{\beta} \right) \right\}$$

Além disso, completando quadrado em relação a uma distribuição normal multivariada, tem-se:

$$P(\boldsymbol{\beta}|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left(\boldsymbol{\beta} - (\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \mathbf{A}_1 \right)^\top (\mathbf{X}_1^\top \mathbf{X}_1) \times \right. \\ \left. \times \left(\boldsymbol{\beta} - (\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \mathbf{A}_1 \right) \right\}.$$

Desta forma, a distribuição completa a posteriori para $\boldsymbol{\beta}$ é normal multivariada.

$$\boldsymbol{\beta}|\dots \sim N \left[(\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \mathbf{A}_1; (\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \sigma_e^2 \right]$$

com $\mathbf{A}_1 = \mathbf{y} - \mathbf{Z}\mathbf{g} - \sum \lambda_k \text{diag}(\mathbf{Z}\boldsymbol{\alpha}_k) \mathbf{X}_2 \boldsymbol{\gamma}_k$.

Distribuição condicional completa a posteriori para \mathbf{g}

A distribuição condicional completa a posteriori para \mathbf{g} é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a \mathbf{g} :

$$P(\mathbf{g}|\dots) \propto |\mathbf{I}\sigma_e^2|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta}) \right\} \exp \left\{ -\frac{1}{2\sigma_g^2} \mathbf{g}^\top \mathbf{I}_g \mathbf{g} \right\} \\ P(\mathbf{g}|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left[(\mathbf{A}_2 - \mathbf{Z}\mathbf{g})^\top (\mathbf{A}_2 - \mathbf{Z}\mathbf{g}) + \frac{\sigma_e^2}{\sigma_g^2} (\mathbf{g}^\top \mathbf{I}_g \mathbf{g}) \right] \right\}. \quad (26)$$

Considerando e desenvolvendo a forma quadrática entre os colchetes e

desconsiderando os termos que não dependem de \mathbf{g} , a expressão (26) pode ser reescrita como:

$$P(\mathbf{g}|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left[-2\mathbf{g}^\top \mathbf{Z}^\top \mathbf{A}_2 + \mathbf{g}^\top \left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right) \mathbf{g} \right] \right\}.$$

Por meio de manipulação algébrica e completando quadrado em relação a uma distribuição normal multivariada, obtém-se:

$$P(\mathbf{g}|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left[\left[\mathbf{g} - \left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right)^{-1} \mathbf{Z}^\top \mathbf{A}_2 \right]^\top \left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right) \times \right. \right. \\ \left. \left. \times \left[\mathbf{g} - \left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right)^{-1} \mathbf{Z}^\top \mathbf{A}_2 \right] \right] \right\}.$$

Assim, a distribuição completa a posteriori para efeito de genótipo é gaussiana multivariada, dada por:

$$\mathbf{g}|\dots \sim N \left[\left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right)^{-1} \mathbf{Z}^\top \mathbf{A}_2; \left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right)^{-1} \sigma_e^2 \right]$$

sendo: $\mathbf{A}_2 = \mathbf{y} - \mathbf{X}_1\beta - \sum \lambda_k \text{diag}(\mathbf{Z}\alpha_k)\mathbf{X}_2\gamma_k$.

Distribuição condicional completa a posteriori para σ_g^2

A distribuição condicional completa a posteriori para σ_g^2 é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a σ_g^2 :

$$P(\sigma_g^2|\dots) \propto (\sigma_g^2)^{-\frac{n_g}{2}} \exp \left\{ -\frac{1}{2(\sigma_g^2)} \mathbf{g}^\top \mathbf{I}_g \mathbf{g} \right\} (\sigma_g^2)^{-1}$$

$$P(\sigma_g^2|\dots) \propto (\sigma_g^2)^{-\left(\frac{n_g}{2}+1\right)} \exp\left\{-\frac{n_g}{2(\sigma_g^2)n_g} \mathbf{g}^\top \mathbf{I}_g \mathbf{g}\right\}.$$

Assim, a distribuição completa a posteriori para variância do efeito do genótipo é uma Qui-quadrado escalada invertida, dada por:

$$\sigma_g^2|\dots \sim Esc - \chi^{-2} \left[n_g; \frac{\mathbf{g}^\top \mathbf{I}_g \mathbf{g}}{n_g} \right].$$

Distribuição condicional completa a posteriori para λ_k

A distribuição condicional completa a posteriori para λ_k é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a λ_k :

$$P(\lambda_k|\dots) \propto \exp\left\{-\frac{1}{2\sigma_e^2}(\mathbf{y} - \theta)^\top (\mathbf{y} - \theta)\right\} P(\lambda_k)$$

$$P(\lambda_k|\dots) \propto \exp\left\{-\frac{1}{2\sigma_e^2}(\mathbf{A}_4 - \lambda_k \phi)^\top (\mathbf{A}_4 - \lambda_k \phi) - \frac{1}{2\sigma_{\lambda_k}^2}(\lambda_k - \mu_{\lambda_k})^\top (\lambda_k - \mu_{\lambda_k})\right\} \quad (27)$$

com $\mu_{\lambda_k} = 0$.

Trabalhando com os termos entre chaves da expressão (27), tem-se:

$$\mathbf{A}_4^\top \mathbf{A}_4 - \lambda_k \mathbf{A}_4^\top \phi - \lambda_k \phi^\top \mathbf{A}_4 - \lambda_k^2 \phi^\top \phi + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2}(\lambda_k^2 - \lambda_k^2 \mu_{\lambda_k} + \mu_{\lambda_k}^2).$$

Considerando e desenvolvendo a forma quadrática e desconsiderando os termos que não dependem de λ_k (estes são absorvidos pela constante de normalização) a expressão (27) acima pode ser reescrita como:

$$P(\lambda_k|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left(-2\lambda_k \phi^\top \mathbf{A}_4 + \lambda_k^2 \left[\phi^\top \phi + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right] \right) \right\}.$$

Por meio de manipulação algébrica e completando quadrado em relação a uma distribuição normal, obtém-se:

$$\propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left(\left[\lambda_k - \left(\phi^\top \phi + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-1} \phi^\top \mathbf{A}_{4k} \right]^\top \left(\phi^\top \phi + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right) \left[\lambda_k - \left(\phi^\top \phi + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-1} \phi^\top \mathbf{A}_{4k} \right] \right) \right\}.$$

Sendo assim, a distribuição completa a posteriori para λ_k é dada por uma normal truncada (N^+):

$$\lambda_k|\dots \sim N^+ \left(\left[\left(\phi_k^\top \phi_k + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-1} \phi_k^\top \mathbf{A}_{4k} \right]; \left(\phi_k^\top \phi_k + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-1} \sigma_e^2 \right)$$

sendo: $\mathbf{A}_{4k} = \mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} - \mathbf{Z}\mathbf{g} - \sum_{k' \neq k}^{t-1} \lambda_{k'} \text{diag}(\mathbf{Z}\boldsymbol{\alpha}_{k'}) \mathbf{X}_2 \boldsymbol{\gamma}_{k'}$ e $\phi_k = \text{diag}(\mathbf{Z}\boldsymbol{\alpha}_k) \mathbf{X}_2 \boldsymbol{\gamma}_k$.

Distribuição condicional completa a posteriori para $\boldsymbol{\alpha}_k$

A distribuição condicional completa a posteriori para $\boldsymbol{\alpha}_k$ é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a $\boldsymbol{\alpha}_k$:

$$P(\boldsymbol{\alpha}_k|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta}) \right\} P(\boldsymbol{\alpha}_k)$$

$$P(\boldsymbol{\alpha}_k|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{A}_4 - \boldsymbol{\Delta}_1 \boldsymbol{\alpha}_k)^\top (\mathbf{A}_4 - \boldsymbol{\Delta}_1 \boldsymbol{\alpha}_k) \right\}. \quad (28)$$

Desenvolvendo os cálculos indicados no numerador da expressão (28) entre chaves, obtém-se:

$$\begin{aligned} (\mathbf{A}_{4k} - \boldsymbol{\Delta}_1 \boldsymbol{\alpha}_k)^\top (\mathbf{A}_{4k} - \boldsymbol{\Delta}_1 \boldsymbol{\alpha}_k) = \\ \mathbf{A}_{4k}^\top \mathbf{A}_{4k} - \mathbf{A}_{4k}^\top \boldsymbol{\Delta}_1 \boldsymbol{\alpha}_k - \boldsymbol{\alpha}_k^\top \boldsymbol{\Delta}_1^\top \mathbf{A}_{4k} + \boldsymbol{\alpha}_k^\top \boldsymbol{\Delta}_1^\top \boldsymbol{\Delta}_1 \boldsymbol{\alpha}_k. \end{aligned}$$

Além disso, considerando e desenvolvendo a forma quadrática e desconsiderando os termos que não dependem de α_k , tem-se a seguinte expressão:

$$P(\boldsymbol{\alpha}_k|\dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left[-2\boldsymbol{\alpha}_k^\top \boldsymbol{\Delta}_1^\top (\mathbf{y} - \mathbf{X}_1 \boldsymbol{\beta} - \mathbf{Z} \mathbf{g}) \right] \right\}$$

sendo assim chega-se que α_k possui distribuição proporcional a von Mises-Fisher (VMF) sendo representada como:

$$\boldsymbol{\alpha}_k|\dots \sim VMF \left(\frac{\lambda_k}{\sigma_e^2}; \boldsymbol{\Delta}_1^\top (\mathbf{y} - \mathbf{X}_1 \boldsymbol{\beta} - \mathbf{Z} \mathbf{g}) \right)$$

com: $\boldsymbol{\Delta}_{1k} = \lambda_k \text{diag}(\mathbf{X}_2 \boldsymbol{\gamma}_k) \mathbf{Z}$.

Distribuição condicional completa a posteriori para $\boldsymbol{\gamma}_k$

A distribuição condicional completa a posteriori para $\boldsymbol{\gamma}_k$ é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a $\boldsymbol{\gamma}_k$:

$$\begin{aligned} P(\boldsymbol{\gamma}_k|\dots) &\propto \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta}) \right\} P(\boldsymbol{\gamma}_k) \\ P(\boldsymbol{\gamma}_k|\dots) &\propto \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{A}_4 - \boldsymbol{\Delta}_2 \boldsymbol{\gamma}_k)^\top (\mathbf{A}_4 - \boldsymbol{\Delta}_2 \boldsymbol{\gamma}_k) \right\}. \quad (29) \end{aligned}$$

De forma análoga à utilizada para vetores genotípicos, desenvolvendo os cálculos indicados na expressão (29) entre chaves, obtém-se:

$$(\mathbf{A}_4 - \mathbf{\Delta}_2 \boldsymbol{\gamma}_k)^\top (\mathbf{A}_4 - \mathbf{\Delta}_2 \boldsymbol{\gamma}_k) = \mathbf{A}_4^\top \mathbf{A}_4 - \mathbf{A}_4^\top \mathbf{\Delta}_2 \boldsymbol{\gamma}_k - \boldsymbol{\gamma}_k^\top \mathbf{\Delta}_2^\top \mathbf{A}_4 + \boldsymbol{\gamma}_k^\top \mathbf{\Delta}_2^\top \mathbf{\Delta}_2 \boldsymbol{\gamma}_k.$$

A partir de manipulações algébricas em que termos não dependentes de $\boldsymbol{\gamma}_k$ são absorvidos pela constante de normalização a expressão é reescrita da seguinte forma:

$$P(\boldsymbol{\gamma}_k | \dots) \propto \exp \left\{ -\frac{1}{2\sigma_e^2} \left[-2\boldsymbol{\gamma}_k^\top \mathbf{\Delta}_2^\top (\mathbf{y} - \mathbf{X}_1 \boldsymbol{\beta} - \mathbf{Z} \mathbf{g}) \right] \right\}.$$

Assim, $\boldsymbol{\gamma}_k$ possui distribuição proporcional a uma von Mises-Fisher (VMF) denotada por:

$$\boldsymbol{\gamma}_k | \dots \sim VMF \left(\frac{\lambda_k}{\sigma_e^2}; \mathbf{\Delta}_2^\top (\mathbf{y} - \mathbf{X}_1 \boldsymbol{\beta} - \mathbf{Z} \mathbf{g}) \right)$$

com média direcional $\mathbf{\Delta}_{2k} = \lambda_k \text{diag}(\mathbf{Z} \boldsymbol{\alpha}_k) \mathbf{X}_2$ e parâmetro de concentração igual a λ_k / σ_e^2 .

Distribuição condicional completa a posteriori para σ_e^2

A distribuição condicional completa a posteriori para σ_e^2 é obtida a partir da equação (24) descartando-se os termos considerados constantes em relação a σ_e^2 :

$$P(\sigma_e^2 | \dots) \propto (\sigma_e^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma_e^2} (\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta}) \right\} (\sigma_e^2)^{-1}.$$

Assim, a distribuição completa a posteriori para a variância do erro do genótipo é uma Qui-quadrado escalada invertida, dada por:

$$\sigma_e^2 | \dots \sim Esc - \chi^{-2} \left[n_e; \frac{(\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta})}{n_e} \right].$$

Distribuição condicional completa a posteriori para t

Dadas as distribuições a priori que correspondem ao parâmetro t tem-se então que a distribuição condicional completa a posteriori para t é dada por:

$$P(t | \dots) \propto P(t | \mu)$$

$$P(t | \dots) \propto \frac{\mu^t e^{-\mu}}{t!}$$

$$t \sim Poisson - truncada(\mu).$$

Distribuição condicional completa a posteriori para μ

Dadas as distribuições a priori que correspondem ao parâmetro μ tem-se então que a distribuição condicional completa a posteriori para μ é dada por:

$$P(\mu | \dots) \propto P(\mu) P(t | \mu)$$

$$P(\mu | \dots) \propto \frac{\mu^t e^{-\mu}}{t!} \frac{\tau^v}{\Gamma(v)} \mu^{v-1} e^{-\tau \mu}$$

$$P(\mu | \dots) \propto \mu^t e^{-2\mu}$$

$$\mu \sim gama(\tau, \mu)$$

sendo: $\tau = 1$ e $v = 1$.

Distribuição condicional completa a posteriori para $\sigma_{\lambda_k}^2$

Para o parâmetro de escala relacionado ao valor singular $\sigma_{\lambda_k}^2$, como já referido, foram estabelecidas três hipóteses (prioris), e para cada uma delas será apresentada a distribuição condicional completa a posteriori correspondente.

Hipótese a: Nesse contexto, é aceita a suposição de que $\sigma_{\lambda_k}^2 \rightarrow \infty$, ou seja, de que o valor da variância é grande, não sendo declarada uma incerteza em relação a $\sigma_{\lambda_k}^2$ por meio de uma distribuição de probabilidade, como feita no caso do vetor de efeitos principais de genótipos (OLIVEIRA et al., 2015). Sob essa hipótese, a distribuição a posteriori para λ_k reduz-se a:

$$\lambda_k | \dots \sim N^+ \left(\left[\left(\phi_k^\top \phi_k \right)^{-1} \phi_k \mathbf{A}_{4k} \right]; \left(\phi_k^\top \phi_k \right)^{-1} \sigma_e^2 \right).$$

Hipótese b: Para a hipótese “b”, a distribuição a priori para $\sigma_{\lambda_k}^2$ é qui-quadrado escalada invertida com parâmetro de escala igual a zero e grau de liberdade $(1 - 2\Delta)$, como em Silva et al. (2015). Então, a partir de cálculos algébricos, obtém-se a distribuição completa a posteriori:

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2(\sigma_{\lambda_k}^2)} (\lambda_k - \mu_{\lambda_k})^\top (\lambda_k - \mu_{\lambda_k}) \right\} P(\sigma_{\lambda_k}^2)$$

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2(\sigma_{\lambda_k}^2)} (\lambda_k - \mu_{\lambda_k})^\top (\lambda_k - \mu_{\lambda_k}) \right\} (\sigma_{\lambda_k}^2)^{(\Delta-1)}$$

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{\frac{1}{2} + (\Delta-1)} \exp \left\{ -\frac{1}{2(\sigma_{\lambda_k}^2)} \lambda_k^2 \right\}$$

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{-\frac{(1-2\Delta)}{2}-1} \exp \left\{ -\frac{(1-2\Delta)}{2(\sigma_{\lambda_k}^2)} \frac{\lambda_k^2}{(1-2\Delta)} \right\}$$

$$\sigma_{\lambda_k}^2 | \dots \sim Esc - \chi^{-2} \left[(1 - 2\Delta), \frac{\lambda_k^2}{(1 - 2\Delta)} \right].$$

Hipótese c: Nessa situação, admite-se que a distribuição a posteriori para $\sigma_{\lambda_k}^2$ é de máxima entropia (APÊNDICE A), sendo, aqui, uma Gama. A distribuição completa a posteriori para $\sigma_{\lambda_k}^2$, a exemplo dos demais parâmetros, é obtida por manipulações algébricas a partir da distribuição conjunta a posteriori da seguinte forma:

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2(\sigma_{\lambda_k}^2)} (\lambda_k - \mu_{\lambda_k})^\top (\lambda_k - \mu_{\lambda_k}) \right\} P(\sigma_{\lambda_k}^2)$$

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2(\sigma_{\lambda_k}^2)} (\lambda_k - \mu_{\lambda_k})^\top (\lambda_k - \mu_{\lambda_k}) \right\} \times \\ \times (\sigma_{\lambda_k}^2)^{-(v_{\lambda_k}-1)} \exp \left\{ -\frac{S_{\lambda_k}^2}{\sigma_{\lambda_k}^2} \right\}$$

$$P(\sigma_{\lambda_k}^2 | \dots) \propto (\sigma_{\lambda_k}^2)^{-\left[\frac{(2v_{\lambda_k}+1)}{2} - 1 \right]} \exp \left\{ -\frac{S_{\lambda_k}^2 + \lambda_k^2}{2(\sigma_{\lambda_k}^2)} \right\}$$

$$\sigma_{\lambda_k}^2 | \dots \sim inv - gama \left(\frac{2v_{\lambda_k} + 1}{2}, \frac{[S_{\lambda_k} + \lambda_k^2]}{2} \right).$$

3.5 Processo de amostragem via MCMC com *Reversible Jump*

3.5.1 Regra de decisão

Como o objetivo é realizar escolha de modelo durante o processo MCMC, optou-se pelo uso do algoritmo saltos reversíveis (*Reversible Jump - RJ*). Desta forma, é possível alternar entre modelos com dimensões diferentes entre os passos MCMC de modo que as propriedades inerentes ao método MCMC não são viola-

das. No contexto, aqui referido, para determinar se a inclusão de eixos principais é necessária no modelo, utiliza-se a regra de decisão $\min[1, a(t, t + 1)]$, semelhante às propostas de Waagepetersen e Sorensen (2001) e Bodin e Sambridge (2009), em que se tem que o determinante do Jacobiano da transformação do modelo t para $t + 1$, da regra de decisão, é igual a um.

Sendo assim a regra de decisão para testar o acréscimo de novo eixo principal ao modelo é dada por:

$$a(t, t + 1) = \frac{\prod_{i=1}^n p(y|t + 1)}{\prod_{i=1}^n p(y|t)} \frac{p(t + 1|\mu)}{p(t|\mu)} \frac{\xi(t, t + 1)}{\xi(t + 1, t)} \quad (30)$$

correspondendo t ao modelo no estado atual e $t + 1$ ao modelo candidato com o acréscimo de um eixo; $p(y|t)$ e $p(y|t + 1)$ são as funções de probabilidade dos dados diante aos parâmetros dos modelos t e $t + 1$, respectivamente; $p(t|\mu)$ e $p(t + 1|\mu)$ são as prioris acerca do modelo com t e $t + 1$ eixos principais, respectivamente; $\xi(t, t + 1)$ e $\xi(t + 1, t)$ correspondem às propostas das distribuições Hastings necessárias para permitir a condição reversível no modelo durante o processo MCMC, sendo $\xi(t, t + 1) = \frac{1}{t+1}pd$ e $\xi(t + 1, t) = p_0$, em que pa , pd e p_0 correspondem às probabilidades a priori de adicionar, excluir ou manter o número de eixos principais no modelo ($pa = pd = p_0 = 1/3$).

Simplificando a expressão (30), tem-se que a regra de decisão para acréscimo de uma dimensão no modelo é:

$$a(t, t + 1) = \frac{\prod_{i=1}^n p(y|t + 1)}{\prod_{i=1}^n p(y|t)} \frac{\mu}{t + 1} \frac{pd}{(t + 1)p_0}$$

sendo $p(y|t)$ e $p(y|t + 1)$ as distribuições condicionais dos dados do modelo atual

(t) e do modelo com acréscimo de uma dimensão ($t + 1$), respectivamente.

Se $a(t, t + 1)$ é maior que uma variável aleatória $u \sim U(0, 1)$, então o eixo candidato é incluído no modelo; caso contrário, o modelo permanece com o número inicial de eixos principais presente no modelo t .

Prosseguindo o processo e considerando realizar o teste entre o modelo no estado atual (t) e o modelo com uma redução de dimensão ($t - 1$), a regra de decisão é dada por:

$$a(t, t - 1) = \frac{\prod_{i=1}^n p(y|t - 1)}{\prod_{i=1}^n p(y|t)} \frac{p(t - 1|\mu) \xi(t, t - 1)}{p(t|\mu) \xi(t - 1, t)} \quad (31)$$

Simplificando a expressão (31), tem-se que, a regra de decisão para excluir uma dimensão no modelo é:

$$a(t, t - 1) = \frac{\prod_{i=1}^n p(y|t)}{\prod_{i=1}^n p(y|t - 1)} \frac{t \ p a \ t}{\mu \ p_0}$$

sendo $p(y|t - 1)$ a distribuição condicional dos dados para o modelo com redução de dimensão ($t - 1$), e $\xi(t, t - 1)$ e $\xi(t - 1, t)$ correspondem às distribuições propostas de Hastings necessárias para permitir o processo reversível no modelo durante a amostragem MCMC, sendo $\xi(t, t - 1) = P_0$ e $\xi(t - 1, t) = Pd/t$.

3.5.2 Algoritmo Gibbs

Nesta subseção, são especificados apenas os passos referentes ao algoritmo Gibbs para todos os parâmetros do modelo, sabendo-se que o processo de amostragem iterativa é uma combinação dos algoritmo de Gibbs e RJMCM. O processo de amostragem Gibbs pode ser conduzido por meio dos seguintes passos:

A1 - Primeiramente, atribuímos-se valores iniciais aos parâmetros do modelo inicial:

$$(\boldsymbol{\beta}^0, \mathbf{g}^0, \lambda_k^0, \boldsymbol{\alpha}_k^0, \boldsymbol{\gamma}_k^0, (\sigma_{\lambda_k}^2)^0, (\sigma_g^2)^0, (\sigma_e^2)^0, \mu^0).$$

A2 - A partir desses valores iniciais, a l -ésima iteração pode ser obtida da seguinte forma:

A. Gera $\boldsymbol{\beta}^l | \mathbf{g}^{l-1}, \lambda_k^{l-1}, (\sigma_{\lambda_k}^2)^{l-1}, \boldsymbol{\alpha}_k^{l-1}, \boldsymbol{\gamma}_k^{l-1}, (\sigma_g^2)^{l-1}, (\sigma_e^2)^{l-1}, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\boldsymbol{\beta} | \dots \sim N \left[(\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \mathbf{A}_1; (\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \sigma_e^2 \right].$$

B. Gera $\mathbf{g}^l | \boldsymbol{\beta}^l, \lambda_k^{l-1}, (\sigma_{\lambda_k}^2)^{l-1}, \boldsymbol{\alpha}_k^{l-1}, \boldsymbol{\gamma}_k^{l-1}, (\sigma_g^2)^{l-1}, (\sigma_e^2)^{l-1}, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\mathbf{g} | \dots \sim N \left[\left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right)^{-1} \mathbf{Z}^\top \mathbf{A}_2; \left(\mathbf{Z}^\top \mathbf{Z} + \frac{\sigma_e^2}{\sigma_g^2} \mathbf{I}_g \right)^{-1} \sigma_e^2 \right]$$

C. Gera $\lambda_k^l | \boldsymbol{\beta}^l, \mathbf{g}^l, (\sigma_{\lambda_k}^2)^{l-1}, \boldsymbol{\alpha}_k^{l-1}, \boldsymbol{\gamma}_k^{l-1}, (\sigma_g^2)^{l-1}, (\sigma_e^2)^{l-1}, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\lambda_k | \dots \sim N^+ \left(\left[\left(\phi_k^\top \phi_k + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-1} \phi_k \mathbf{A}_{4k} \right]; \left(\phi_k^\top \phi_k + \frac{\sigma_e^2}{\sigma_{\lambda_k}^2} \right)^{-1} \sigma_e^2 \right).$$

Com relação ao passo “D”, são especificadas três hipóteses referentes à componente de variância relacionada ao valor singular. Para o esquema de amostragem Gibbs, apenas uma das hipóteses será usada para ilustrar o processo de amostragem.

D. Gera $(\sigma_{\lambda_k}^2)^l | \boldsymbol{\beta}^l, \mathbf{g}^l, \lambda_k^l, \boldsymbol{\alpha}_k^l, \boldsymbol{\gamma}_k^l, (\sigma_e^2)^{l-1}, (\sigma_g^2)^{l-1}, \mu^{l-1}$ a partir da distribuição

condicional a posteriori:

$$\sigma_{\lambda_k}^2 | \dots \sim \text{inv-gama} \left(\frac{2\nu_{\lambda_k} + 1}{2}; \frac{[S_{\lambda_k} + \lambda_k^2]}{2} \right).$$

E. Gera $\alpha_k^l | \beta^l, \mathbf{g}^l, \lambda_k^l, (\sigma_{\lambda_k}^2)^l, \gamma_k^{l-1}, (\sigma_g^2)^{l-1}, (\sigma_e^2)^{l-1}, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\alpha_k | \dots \sim \text{VMF} \left(\frac{\lambda_k}{\sigma_e^2}; \mathbf{\Delta}_1^\top (\mathbf{y} - \mathbf{X}_1 \beta - \mathbf{Z} \mathbf{g}) \right).$$

F. Gera $\gamma_k^l | \beta^l, \mathbf{g}^l, \lambda_k^l, (\sigma_{\lambda_k}^2)^l, \alpha_k^l, (\sigma_g^2)^{l-1}, (\sigma_e^2)^{l-1}, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\gamma_k | \dots \sim \text{VMF} \left(\frac{\lambda_k}{\sigma_e^2}; \mathbf{\Delta}_2^\top (\mathbf{y} - \mathbf{X}_1 \beta - \mathbf{Z} \mathbf{g}) \right).$$

G. Gera $(\sigma_g^2)^l | \beta^l, \mathbf{g}^l, \lambda_k^l, (\sigma_{\lambda_k}^2)^l, \alpha_k^l, \gamma_k^l, (\sigma_e^2)^{l-1}, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\sigma_g^2 | \dots \sim \text{Esc} - \chi^{-2} \left[n_g; \frac{\mathbf{g}^\top \mathbf{I}_g \mathbf{g}}{n_g} \right].$$

H. Gera $(\sigma_e^2)^l | \beta^l, \mathbf{g}^l, \lambda_k^l, (\sigma_{\lambda_k}^2)^l, \alpha_k^l, \gamma_k^l, (\sigma_g^2)^l, \mu^{l-1}$ a partir da distribuição condicional a posteriori:

$$\sigma_e^2 | \dots \sim \text{Esc} - \chi^{-2} \left[n_e; \frac{(\mathbf{y} - \boldsymbol{\theta})^\top (\mathbf{y} - \boldsymbol{\theta})}{n_e} \right].$$

I. Gera $\mu^l | \beta^l, \mathbf{g}^l, \lambda_t^l, \alpha_t^l, \gamma_t^l, (\sigma_{\lambda_t}^2)^l, (\sigma_g^2)^l, (\sigma_e^2)^l$ a partir da distribuição condicio-

nal a posteriori:

$$\mu \sim \text{gama}(\tau; \mu).$$

Após ter realizado o processo amostral dos parâmetros estimáveis do modelo (t) da l -ésima iteração, amostram-se os parâmetros adicionais referentes ao modelo concorrente ($t + 1$).

A dimensão do modelo é dada pelo número de termos multiplicativos presentes nele ($0, 1, 2, \dots, \text{posto}(GEI)$). Para cada proposta de modelo, essa dimensão está diretamente relacionada aos passos C, D, E e F do processo de amostragem Gibbs. Para o modelo sem a presença da parte multiplicativa $t = 1$, os passos C, D, E e F , não são incluídos no processo de amostragem; apenas para $t > 1$, os passos são incluídos no processo de estimação, sempre em relação ao número de termos presente para cada modelo candidato.

3.5.3 Algoritmo RJMCMC

Na seção anterior, foram descritos o processo de amostragem para os parâmetros dos modelos candidatos pelo amostrador de Gibbs, a regra de decisão para seleção do modelo mais provável e os principais aspectos relacionados à abordagem aqui proposta. Nesta seção, será especificado o algoritmo de forma esquemática para alternância entre modelos.

Dado o estado atual da cadeia (t, θ_t), o processo amostral de RJMCMC é conduzido da seguinte forma:

Inicia-se o processo Gibbs para amostragem dos parâmetros contidos no modelo para t e, então, todos os passos presentes no algoritmo são executados. No caso de $t = 1$, os passos C, D, E e F são excluídos do algoritmo, permanecendo apenas os efeitos principais do modelo para o processo de amostragem.

i) Movimento de inclusão:

- 1 Propor o acréscimo de um termo bilinear no modelo $(t + 1)$.
- 2 Calcular a probabilidade de aceitar a inclusão do termo bilinear no modelo:

$$a(t, t + 1) = \frac{\prod_{i=1}^n p(y|t + 1)}{\prod_{i=1}^n p(y|t)} \frac{\mu}{t + 1} \frac{pd}{(t + 1)p_0}.$$

- 3 Amostrar um número aleatório de uma distribuição uniforme $u \sim U(0, 1)$.
- 4 Se $u < \min(1, a(t, t + 1))$, aceita-se o estado proposto; caso contrário, o modelo permanece com a dimensão inicial.

ii) Movimento de exclusão:

- 1 Propor a remoção de um termo bilinear no modelo $(t - 1)$.
- 2 Calcular a probabilidade de aceitar a exclusão de um termo bilinear no modelo:

$$a(t, t - 1) = \frac{\prod_{i=1}^n p(y|t)}{\prod_{i=1}^n p(y|t - 1)} \frac{t}{\mu} \frac{pa}{p_0}.$$

- 3 Amostrar um número aleatório de uma distribuição uniforme $u \sim U(0, 1)$.
- 4 Se $u < \min(1, a(t, t + 1))$, aceita-se o estado proposto; caso contrário, o modelo permanece com a dimensão inicial.

Repetir os passos especificados em i) e ii) durante o processo iterativo.

3.6 Capacidade preditiva do modelo

Para avaliar a capacidade preditiva foi realizado um processo de desbalanceamento nos dados de forma que dentro de cada ambiente foram estabelecidas perdas aleatórias de genótipos (perda total do genótipo). Foi estabelecido apenas um nível de perda de 10% dos dados. Assim, a amostra original foi dividida aleatoriamente em 10 subamostras (*10-fold*) de tamanho igual. Em cada fold os dados foram divididos em dados de treinamento e validação. As perdas simuladas foram realizadas nas células da tabela GEI onde algumas informações de genótipos (linhas) foram totalmente retiradas de ambientes específicos (colunas), mas mantendo todos os ambientes no conjunto de dados, como realizado por Nuvunga et al. (2015).

Esses dados foram usados para avaliar a capacidade preditiva dos modelos BAMMI, BAMMIS e BMMIE, em todas possíveis configurações de dimensão de modelo, como também suas versões utilizando o algoritmo *Reversible Jump*.

A capacidade preditiva dos modelos será medida utilizando os critérios da PRESS (*Predicted Residual Sum of Squares*) e da correlação entre os valores fenotípicos preditos (\hat{y}_{ij}) e observado (y_{ij}). Essa verificação será dada nos cenários envolvendo as diferentes hipóteses sobre a variância relacionada ao valor singular λ_k .

A expressão da correlação é dada por:

$$Cor = \frac{\sum_{i=1}^n (\hat{y}_{ij} - \bar{\hat{y}}_{ij}) (y_{ij} - \bar{y}_{ij})}{\sqrt{\sum_{i=1}^n (\hat{y}_{ij} - \bar{\hat{y}}_{ij})^2} \sqrt{\sum_{i=1}^n (y_{ij} - \bar{y}_{ij})^2}}.$$

A expressão da PRESS é dada por:

$$PRESS = \frac{1}{n} \sum_{j=1}^n (y_{ij} - \hat{y}_{ij})^2.$$

3.7 Diagnóstico das cadeias de Markov e inferência a posteriori

O modelo escolhido para representar o conjunto de dados é aquele com a maior frequência de seleção, com base na regra de decisão. A convergência das cadeias produzidas, de acordo com os passos descritos acima, foi realizada pelo método de Raftery e Lewis (1992) e pelo critério de Heidelberger e Welch (1983), implementados no pacote BOA (Bayesian Output Analysis). Para definir o número de iterações necessário na implementação do amostrador de Gibbs utilizou-se uma amostra piloto. Essa amostra foi avaliada por meio do critério de Raftery e Lewis (1992) que indicou a simulação de cadeias MCMC com tamanho 160.000 com o descarte das primeiras 8.000 observações (*burn-in*) e a coleta dos dados realizando-se saltos a cada 20 observações (*thinning*). Ao fim desse processo foi selecionado um total de 8.000 observações para o processo de inferência.

Os parâmetros lineares do modelo selecionado serão estimados pelas médias a posteriori, com a exceção dos vetores singulares que devem ser estimados a partir de uma correção para preservar as restrições de ortonormalidade entre os vetores (LIU, 2001).

Além disso, a incerteza em relação à estimação dos parâmetros será quantificada por meio de regiões de credibilidade com máxima densidade a posteriori. Regiões bivariadas serão incorporadas ao biplot-AMMI utilizando o método descrito por Hu e Yang (2013), implementado no pacote *distfree.cr* do Software R, para o estudo de estabilidade e adaptabilidade. Todas as análises foram realizadas utilizando o software estatístico R (R CORE TEAM, 2017)

4 RESULTADOS

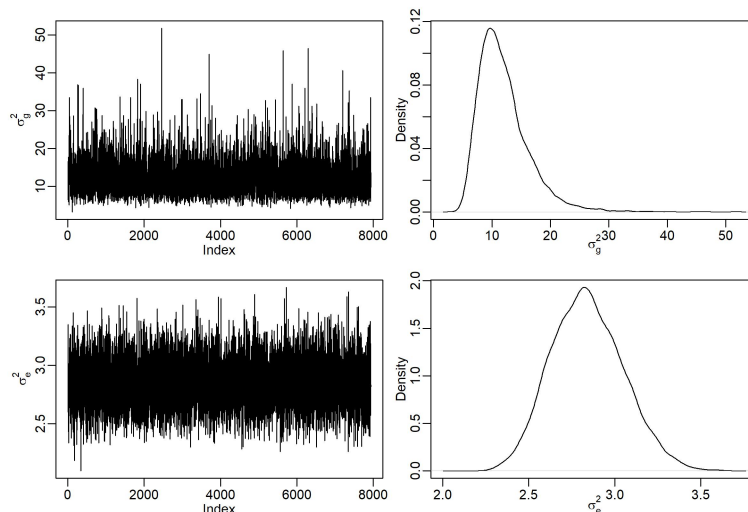
Nessa seção serão apresentados os resultados obtidos a partir dos diferentes critérios utilizados para a seleção de modelos quando se utiliza o método bayesiano, bem como os resultados da proposta aqui apresentada, ou seja, o *Reversible Jump*.

4.1 Convergência e efeitos principais genótipos

Conforme relatado em métodos, a amostragem para os modelos deste trabalho foi realizada por meio do algoritmo Gibbs, no primeiro caso em que o espaço do modelo é fixo, e pelo. Os resultados obtidos pelos critérios de Heidelberg e Welch (1983) e Raftery e Lewis (1992) indicaram boas propriedades de convergência para todos os parâmetros com fator de dependência sempre menor que cinco ($I < 5$). Além disso, todos os parâmetros passaram no teste de estacionariedade, indicando que a convergência foi alcançada.

A Figura 1 mostra os traços das componentes de variância dos modelos em análise, para o modelo BAMMI-RJ, que como pode ser observado apresenta um padrão de dispersão aparentemente uniforme em torno de um valor médio, corroborando os resultados dos critérios de convergências utilizados. Esse padrão também foi verificado nos modelos BMMIE-RJ e BMMIS-RJ que estão apresentados em Apêndice C (Figura 21 e 22).

Figura 1 Traço e densidades das cadeias para os comentários do modelo BAMMI-RJ.



Fonte: Do autor (2019)

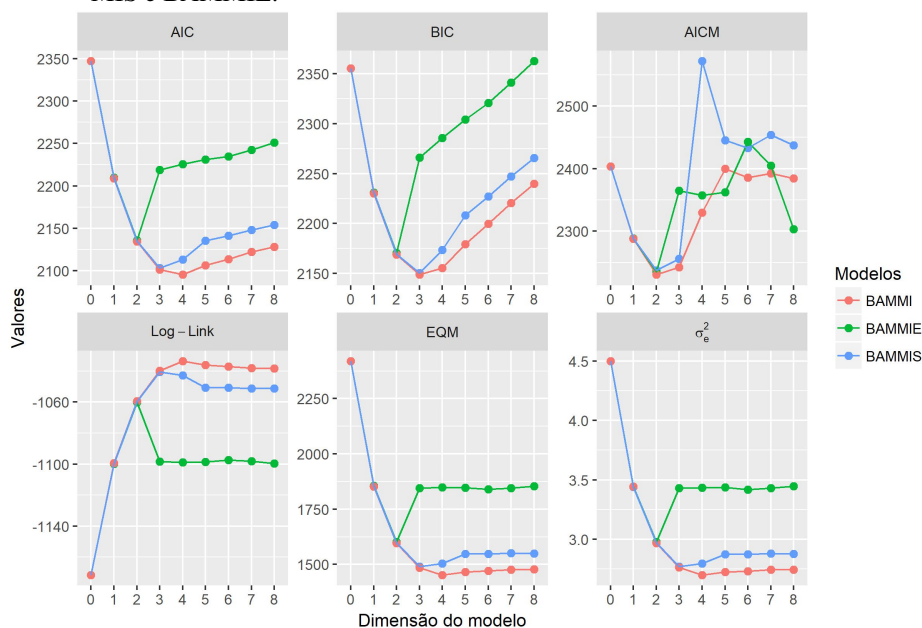
Estimativas para efeitos principais de genótipos juntamente com regiões a 95% de credibilidade para cada modelo selecionado pelos critérios de informações são apresentadas no Apêndice C (Figuras 18, 19 e 20). Os genótipos estão ranqueados da esquerda para direita e sobreposições indicam efeitos semelhantes.

4.2 Seleção de modelo utilizando critério de informação e efeitos principais de genótipos

Como especificado em métodos, foram calculados valores para os critérios de informação, em cada membro da família AMMI, considerando as três abordagens AMMI-bayesiana em função da densidade a priori assumida para o valor singular. Os resultados mostrados na Figura 2, permitiram verificar que o critério que o AICM selecionou o mesmo modelo (segundo o número de termos bilineares) para as três versões AMMI-Bayesiano indicando modelos com dois termos bilineares-AMMI2. Para o critério AIC, os modelos selecionados foram BAM-

MIE2, BAMMIS3 e BAMMI4, ao passo que o BIC selecionou, respectivamente, os modelos Bammie2, Bammis3 e Bammis3.

Figura 2 Valores dos critérios de informação AIC, AICM e BIC, bem como para log-verossimilhança (Log-Lik), Erro quadrático médio (EQM) e variância residual (σ_e^2) em função da dimensão do modelo, para família de modelos BAMMI, Bammis e Bammie.



Fonte: Do autor (2019)

Além disso, a Figura 2 oferece uma descrição do comportamento dos modelos no que se refere aos valores da log-verossimilhança (Log-Lik), erro quadrático médio (EQM) e da variância residual (σ_e^2) para a família de modelos AMMI. No que se refere ao Log-Lik, quanto maior o valor melhor o desempenho do modelo, ao passo que para EQM e σ_e^2 , destacam-se aqueles com menores valores.

Destacamos também, na Tabela 1, outro aspecto relevante no que se refere à validação de modelos, que é a soma de quadrado da GEI acumulada em relação ao número de termos multiplicativos retidos no modelo. A soma de quadrado verdadeira (da matriz GEI simulada) foi de 259,6064. Um fato a ser observado

é que para nenhum membro da família AMMI, o modelo BMMIE superestima o valor da variância total da matriz simulada, embora para todos os membros as estimativas são sempre menores em relação ao BMMIS e BAMMI.

Tabela 1 Valores singulares estimados para os modelos AMMI conforme o número de eixos incorporados ao modelo, e sua soma de quadrado da interação acumulada (Algoritmo Gibbs).

BAMMI									
Dim.	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	SQ(GEI)
1	12,97								168,24
2	13,13	8,63							246,84
3	13,17	8,22	6,12						278,36
4	13,20	8,25	5,38	3,83					285,96
5	13,21	8,22	5,37	2,66	1,77				281,07
6	13,21	8,25	5,37	2,67	1,24	0,75			280,51
7	13,19	8,22	5,32	2,64	1,23	0,59	0,33		278,73
8	13,19	8,22	5,33	2,63	1,23	0,57	0,27	0,14	278,79
BMMIS									
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	SQ(GEI)
1	12,87								165,70
2	13,05	8,50							242,76
3	13,10	8,08	5,91						271,89
4	13,09	8,09	4,18	2,45					260,49
5	13,07	7,98	3,85	0,56	0,11				249,62
6	13,08	7,99	3,87	0,55	0,05	0,01			250,31
7	13,08	7,99	3,78	0,55	0,05	0,001	0,001		249,39
8	13,07	7,99	3,83	0,52	0,05	0,001	0,001	0,001	249,66
BMMIE									
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	SQ(GEI)
1	12,65								160,09
2	12,87	8,23							233,31
3	12,68	0,27	0,001						160,90
4	12,66	0,23	0,001	0,001					160,29
5	12,62	0,27	0,001	0,001	0				159,34
6	12,67	0,55	0,002	0,001	0	0			160,73
7	12,65	0,39	0,001	0,001	0	0	0		160,17
8	12,65	0,05	0,001	0,001	0	0	0	0	160,15

Além disso, a Tabela (1) apresenta a média a posteriori dos valores singulares para os modelos condicionais ao número de termos bilineares ajustados para BAMMI, BAMMI-S e BAMMI-E. É possível observar efeito de encolhimento acentuado do modelo BAMMI-E em relação aos demais a partir do ajuste do terceiro termo multiplicativo. Além disso, observa-se também encolhimento de estimativas do BAMMI-S em relação ao BAMMI, a partir do modelo AMMI5.

Relacionando os resultados observados na Figura 2 com os valores da Tabela 1, observa-se que para o critério AICM os modelos selecionados em ambas as análises não oferecem valores de soma de quadrados acumulada maior que a soma total simulada. Para os critérios AIC e BIC, nos modelos selecionados para BAMMIS e BAMMI, as estimativas da soma de quadrado recuperada pelo modelo é maior em relação à soma de quadrado total, como já especificada.

Para comparação apresentam-se na Tabela 2, os valores “verdadeiros” para os λ_k que foram obtidos pela decomposição em valores singulares da matriz de interação GEI simulada, bem como a soma de quadrado acumulada de acordo com o número de termos bilineares.

Tabela 2 Valores singulares obtidos por meio de decomposição de valores singulares da matriz de interação simulada, assim como o percentual referente a cada eixo principal como o acumulado em cada.

Dimensão	λ_k	λ_k^2	$\lambda_k^2(\%)$	Acum.	Acum.(%)
1	11,82	139,769	0,538	139,769	0,54
2	6,56	43,074	0,166	182,843	0,70
3	5,03	25,348	0,098	208,192	0,80
4	4,17	17,354	0,067	225,546	0,87
5	3,70	13,66	0,053	239,206	0,92
6	3,25	10,556	0,041	249,762	0,96
7	2,63	6,893	0,027	256,655	0,99

8	1,72	2,949	0,011	259,604	1,00
---	------	-------	-------	---------	------

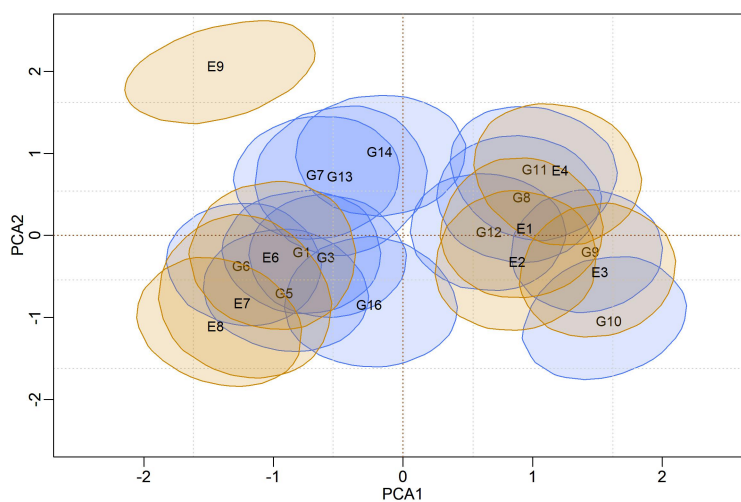
Outro fato a ser destacado é que o modelo BMMIE selecionado por todos os critérios de informação é aquele com dois termos multiplicativos e, também, o que apresentou soma acumulada menos discrepante em relação aos dois primeiros termos (Tabela 2). Por outro lado, em relação aos modelos completos o BAMMIE é o que mais subestima a soma de quadrado total, enquanto o BAMMI seria aquele que mais superestima. O modelo BMMIS, como pode ser visto, oferece valor mais próximo em relação à soma de quadrados total “verdadeira”, quando se considera modelos de maior dimensionalidade.

4.2.1 Regiões de Credibilidades para os modelos AMMI-Bayesiano

Nesta seção são apresentadas as regiões de credibilidade para os modelos BAMMI, BMMIE e BMMIS, para o biplot AMMI2 (dois primeiros eixos singulares). O modelo utilizado para a representação foi o completo (AMMI8) para BAMMI e BMMIS e para BMMIE (AMMI2).

Na Figura 3 são apresentados apenas os escores genotípicos e ambientais cujas regiões de credibilidade (a 95%) não incluem a origem (0,0), ou seja, genótipos e ambientes cujo efeito da interação difere de zero. Esses são os genótipos e ambientes que contribuiriam significativamente para o efeito da interação. O subgrupo de genótipos estáveis, por sua vez, é dado pelos genótipos cujas regiões para os escores englobam a origem G2, G4, G15, G17, G18, G19, G20 e não foram representados no biplot para simplificar as interpretações. Em relação aos ambientes somente o ambiente E5 foi considerado estável.

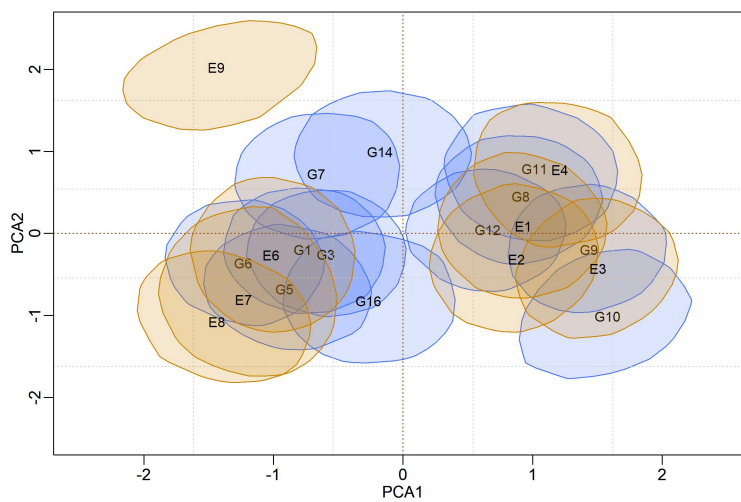
Figura 3 Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMI-8, cujas regiões não incluem a origem (0,0) usando algoritmo Gibbs.



Fonte: Do autor (2019)

De forma semelhante, o biplot para BAMMIS em relação aos dois primeiros eixos principais é apresentado na Figura 4. Utilizando a mesma interpretação, o subgrupo de genótipos G2, G13, G4, G15, G17, G18, G19 e G20 é considerado estável e não foram representados no biplot. Assim como aconteceu para o modelo BAMMI apenas o E5 é considerado estável.

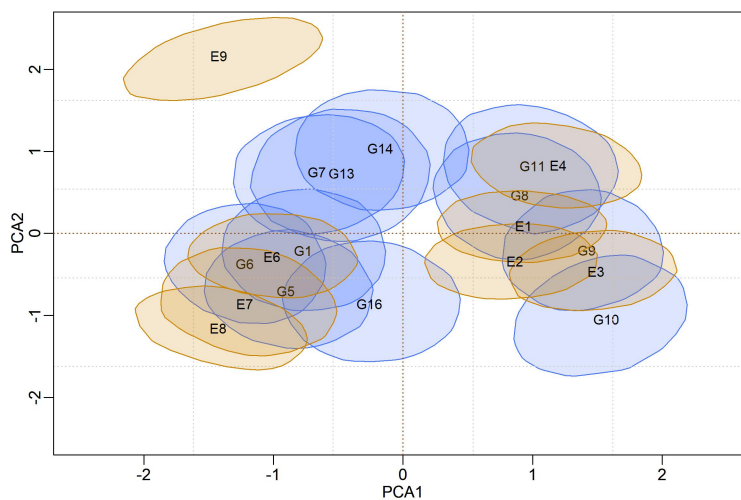
Figura 4 Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMIS-8, cujas regiões não incluem a origem (0,0) usando algoritmo Gibbs.



Fonte: Do autor (2019)

Por fim, a Figura 5 apresenta o biplot para o modelo BAMMIE sendo o subgrupo de genótipos estáveis dado por G2, G3, G4, G12, G15, G17, G18, G19, G20. Nesta abordagem também, o único ambiente considerado estável é o E5.

Figura 5 Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMIE-2, cujas regiões não incluem a origem (0,0) usando algoritmo Gibbs.



Fonte: Do autor (2019)

Além disso, subgrupo de genótipos e/ou ambientes que possuem efeitos semelhantes com relação à GEI podem ser identificados a partir de sobreposições entre as regiões de credibilidades bivariadas. Adaptabilidade de genótipo (ou subgrupo de genótipos) a ambientes (ou subgrupo de ambientes) também pode ser observada por meio de sobreposições, ou ainda observando genótipos e ambientes cujos escores pertencem ao mesmo quadrante.

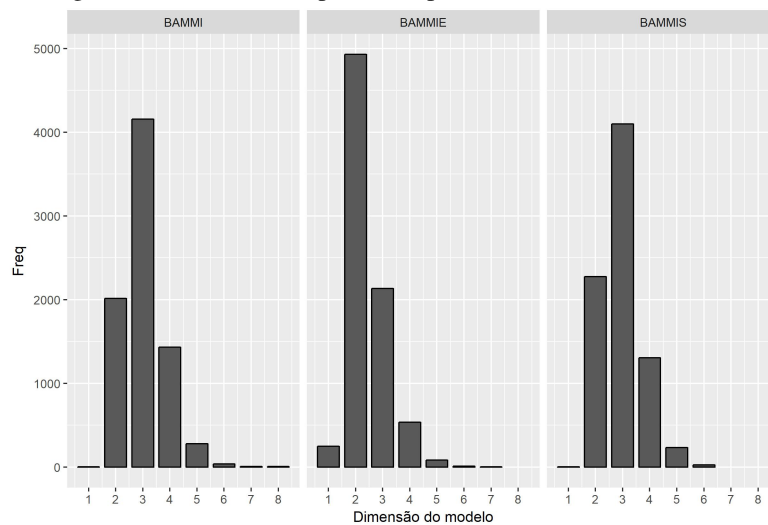
4.3 Ajuste do modelo AMMI pelo método *Reversible Jump*

As estimativas a posteriori quando utilizado o *Reversible Jump* foram obtidas sob duas perspectivas: com base no modelo mais provável e pelo modelo marginal. O modelo mais provável é aquele que considera apenas as observações contidas no modelo condicional (n° de eixos do modelo com maior frequência de “visitas”). Já o modelo marginal considera todas as observações MCMC dos mo-

delos da família AMMI que contem os parâmetros estimados presentes no modelo mais provável, ou seja, considera as estimativas dos modelos aninhados em que o número de termos bilineares é igual ou superior ao do modelo mais provável.

Na Figura 6 apresentam-se as frequências com que cada modelo foi visitado, sendo que as respectivas frequências foram utilizadas como critério para selecionar o modelo vencedor (mais provável). Segundo esse critério, seriam selecionados para BAMMI, BMMIE e BMMIS, os modelos AMMI3, AMMI2 e AMMI3, respectivamente. Notam-se sempre diferenças expressivas entre a frequência do modelo vencedor em relação aos demais para cada versão AMMI.

Figura 6 Probabilidade a posteriori para dimensão (t) do modelo.



Fonte: Do autor (2019)

Na Tabela 3 observa-se que os modelos apresentaram estimativas muito próximas para a média dos dois primeiros valores singulares. Além disso, BAMMI e BMMIS obtiveram valores médios semelhantes também em relação a λ_3 . Ainda em relação a λ_3 observa-se diferença mais expressiva para o BMMIE, embora os intervalos HPD se sobreponham a 95% de probabilidade.

Tabela 3 Valores singulares estimados para o modelo AMMI referentes às três hipóteses realizadas sob a interação, utilizando o algoritmo *Reversible Jump*.

Modelo	Valor Singular	Média	DP	LI	LS
BAMMI	λ_1	13,1993	1,0088	11,3004	15,2613
	λ_2	8,2067	1,0520	6,1417	10,2358
	λ_3	5,5862	1,1914	3,2555	7,8862
BAMMIE	λ_1	12,8496	1,0481	10,8374	14,9333
	λ_2	7,2263	2,0593	5,2887	9,8850
	λ_3	0,4213	1,3676	< 0,001	4,5682
BAMMIS	λ_1	13,0974	1,0059	11,0445	15,0142
	λ_2	8,0646	1,1067	5,8534	10,2091
	λ_3	5,2633	1,4088	2,4312	7,7224

A soma de quadrado referente aos eixos dos modelos selecionados pelo método *Reversible Jump* em relação aos modelos BAMMI, BAMMIE e BAMMIS são respectivamente: 272,7710; 217,3316 e 264,2804, respectivamente. Sendo assim é possível observar, que para os modelos mais frequente em BAMMI e BAMMIS superestimaram a soma de quadrado total da GEI (simulada), enquanto BAMMIE apresentou um valor relativamente menor, embora bem mais próximo da soma de quadrado acumulada pelos dois primeiros eixos como pode ser visto na Tabela 2.

É notável o efeito *shrinkage* acentuado para a média a posteriori de λ_3 no modelo BAMMIE, em relação aos demais, bem como em relação aos valores simulados (Tabela 2), o que indica que as estimativas para os demais parâmetros devem ser encolhidas para o valor zero. Efeito de encolhimento acentuado para as estimativas dos valores singulares nos demais modelos, até o terceiro eixo, não foram observados.

Estimativas pontuais a posteriori juntamente com regiões de credibilidade HPD para variâncias residuais e genotípicas dos modelos mais prováveis segundo

o método *Reversible Jump* (condicionais aos modelos vencedores), bem como para modelos marginais são apresentadas na Tabela 4. Não foram observadas grandes discrepâncias entre as variâncias do modelo condicional e modelo marginal, embora os valores tenham sido sempre menores para variância marginal, e mesmo entre as variâncias das diferentes versões AMMI abordadas não se observaram diferenças acentuadas entre os valores das estimativas.

Tabela 4 Médias a posteriori, desvio padrão (Dp) e limites inferiores (LI) e superiores (LS) de regiões a 95% de credibilidade das variâncias (Var.) genotípicas (σ_g^2) e residuais (σ_e^2) para os modelos condicionais (mais prováveis pelo método *Reversible Jump*) e modelos marginais referente as três versões AMMI bayesiano analisadas.

	Modelo	Var.	Média	Dp.	LI	LS
BAMMI	Marginal	σ_g^2	11,973	4,330	5,258	20,228
		σ_e^2	2,841	0,209	2,454	3,257
	Condicional	σ_g^2	12,028	4,381	5,367	20,391
		σ_e^2	2,810	0,182	2,464	3,161
BAMMIS	Marginal	σ_g^2	12,026	4,386	5,205	20,626
		σ_e^2	2,869	0,210	2,470	3,277
	Condicional	σ_g^2	12,043	4,525	5,205	20,980
		σ_e^2	2,834	0,187	2,473	3,188
BAMMIE	Marginal	σ_g^2	12,028	4,514	5,181	20,735
		σ_e^2	3,054	0,223	2,658	3,530
	Condicional	σ_g^2	12,032	4,567	5,181	20,783
		σ_e^2	3,050	0,209	2,676	3,486

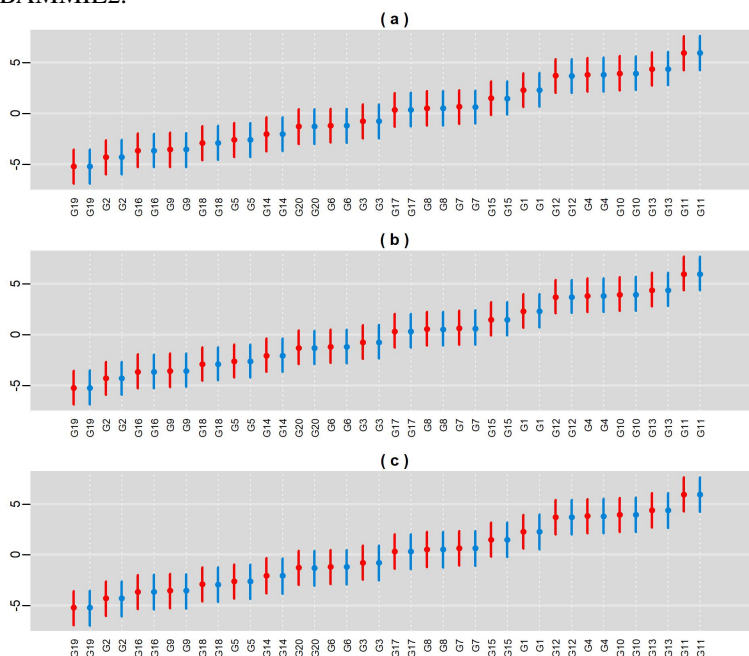
Comparando com os valores dos modelos selecionados pelos critérios de informação aqui utilizados verifica-se também, que não foram observadas grandes discrepâncias no que se refere a variâncias genotípicas e residuais (Apêndice B - Tabelas 5, 6 e 7).

Na Figura 7 são apresentadas médias a posteriori e os intervalos HPD a 95% de credibilidade, para o efeito genotípico referente aos modelos ajustados

pelo método *Reversible Jump*. Foram considerados o modelo condicional ou mais provável (cor azul) e o modelo marginal (cor vermelha). Em ambas as situações o padrão na classificação dos efeitos principais genotípicos foram praticamente o mesmo, para os três modelos. Sobreposição entre os intervalos indicam efeitos semelhantes.

Em relação à produtividade média pode-se selecionar o mesmo subgrupo de genótipos (mais produtivos) G1, G12, G4, G10, G13, G11 para as três versões AMMI. Por outro lado, o subgrupo G19, G2, G16, G9, G18, G5, G14 possui média significativamente menor que a média geral (efeito negativo) não sendo interessante em termos de produtividade. Por fim, têm-se aqueles genótipos cujas regiões HPD incluem a origem e, portanto os efeitos não diferem significativamente de zero. As diferenças para valores das estimativas de efeito principal entre os modelos são praticamente imperceptíveis ao se observar os gráficos. Para melhor discriminar esses resultados apresentam-se nas (Tabelas 8, 9 e 10) com estatísticas pontuais e intervalares (Apêndice B).

Figura 7 Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional ao modelo mais provável pelo Reversible Jump (cor azul) e para o modelo marginal (cor vermelha): a) BAMMI3; b) BAMMIS3; c) Bammie2.



Fonte: Do autor (2019)

4.3.1 Regiões de credibilidades bivariadas em *Reversible Jump*

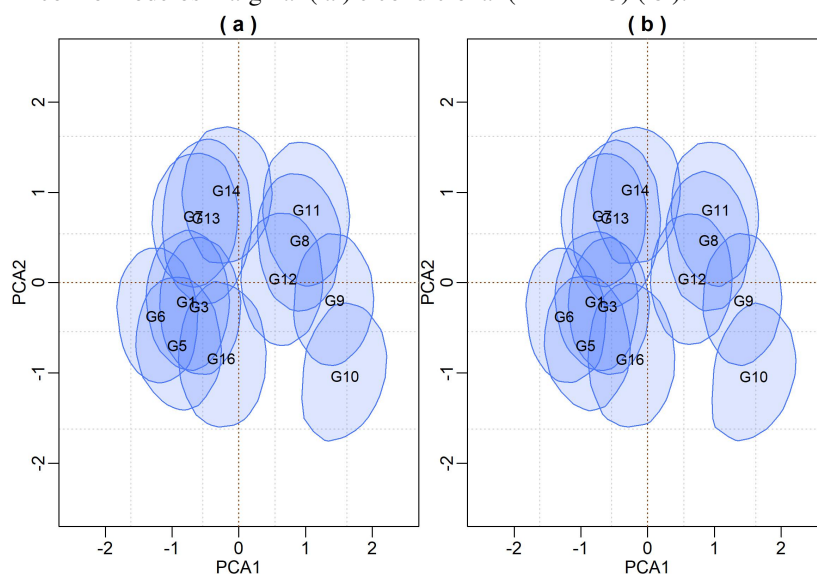
Nesta seção serão apresentados os resultados dos modelos AMMI Bayesiano utilizando o algoritmo *Resersible Jump*, para as três hipóteses aqui propostas na seção (4.3).

4.3.1.1 Modelo BAMMI-RJ

Na Figura 8 são apresentados os biplots, juntamente com as regiões de credibilidades a posteriori marginais (Figura 8 - a) e condicionais ao modelo BAMMI3 (Figura 8 - b), respectivamente. As regiões de credibilidade marginais foram ob-

tidas, considerando os escores dos dois primeiros eixos principais em relação a todos os modelos “visitados” com exceção ao modelo com apenas um termo bilinear (BAMMI1). As regiões bivariadas condicionais foram estabelecidas considerando apenas o modelo mais “visitado” (BAMMI3). Como pode-se observar os dois biplots (condicional e marginal) possuem o mesmo padrão.

Figura 8 Regiões bivariadas a 95% de credibilidades para o modelo BAMMI, construídas com o modelos marginal (a) e condicional (BAMMI3) (b).

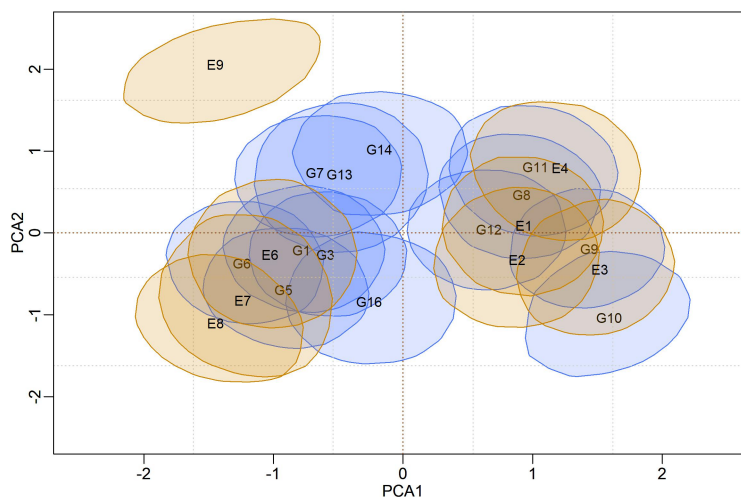


Fonte: Do autor (2019)

Na Figura 9 é exibido o biplot marginal, agora com regiões de credibilidade para escores genotípicos e ambientais (a 95% de probabilidade), que não contêm a origem (0,0). Como já destacado, genótipos e ambientes associados a essas regiões contribuem significativamente para o efeito da interação e frequentemente são referenciados como não estáveis. Nesse sentido, os genótipos do subgrupo G2, G4, G15, G17, G18, G19, G20 podem ser interpretados como estáveis (regiões englobam a origem) e os demais, como já enfatizado, são não estáveis. A mesma análise pode ser realizada em relação aos ambientes, observando que

apenas a região de credibilidade para o escore E5 engloba a origem indicando que o respectivo ambiente pode ser considerado estável.

Figura 9 Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMI cujas regiões não contêm a origem (0,0) (modelo marginal).



Fonte: Do autor (2019)

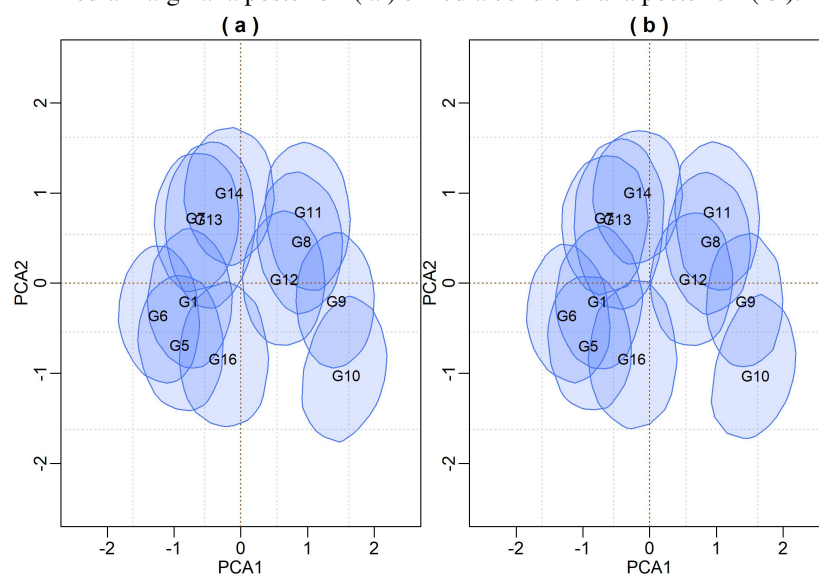
Além disso, é possível identificar subgrupos semelhantes de genótipos e ambientes com relação ao efeito da interação (subgrupos homogêneos). Embora com algumas sobreposições que dificultam a observação de grupos separáveis poderia se pensar em cinco subgrupos homogêneos de genótipos G10, G8, G11, G12, G7, G13, G14 e por último o subgrupo G2, G4, G15, G17, G18, G19, G20 de genótipos considerados estáveis.

Os ambientes, por outro lado, oferecem quatro subgrupos homogêneos distintos que são: E9, E6, E7, E8, E1, E2, E3, E4 e E5. Adaptabilidade de genótipos a ambientes específicos também pode ser sugerida a partir da representação biplot observando regiões que estão no mesmo quadrante e/ou que possuem sobreposições expressivas, como já ressaltado nesse texto.

4.3.1.2 Modelo BAMMIS-RJ

Assim como no modelo BAMMI, na Figura 10 estão dispostas as regiões de credibilidades para dos escores genotípicos considerando os escores marginais a posteriori (Figura 10 - a) e escores condicionais a posteriori para o modelo BAMMIS (Figura 10 - b). Aqui também observaram-se padrões semelhantes nas duas representações gráficas, com leves discrepâncias.

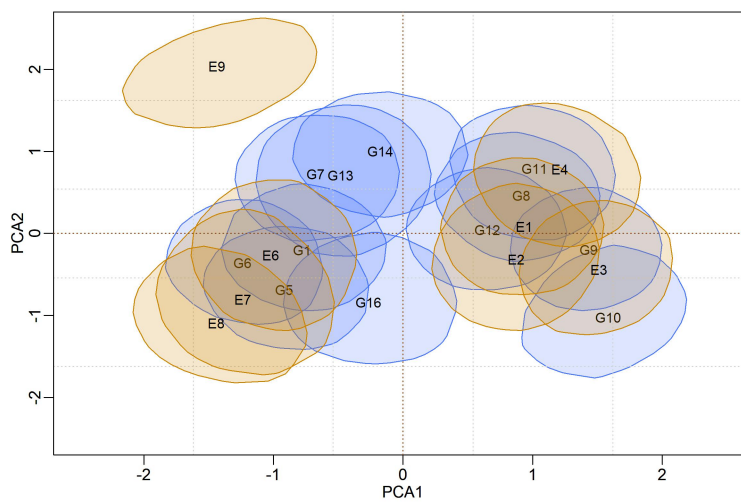
Figura 10 Regiões de credibilidades para o modelo BAMMIS, construídas com base na média marginal a posteriori (a) e média condicional a posteriori (b).



Fonte: Do autor (2019)

O biplot incluindo escores ambientais para o modelo marginal é apresentado na Figura 11. Interpretações semelhantes àquelas realizadas para a análise BAMMI podem ser feitas, observando sobreposições entre as regiões de credibilidade para os escores. Uma diferença em relação aos dois biplots é que G3 na análise BAMMIS é interpretado como estável ao passo que para BAMMI é não estável.

Figura 11 Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMIS cujas regiões não contêm a origem (0,0).

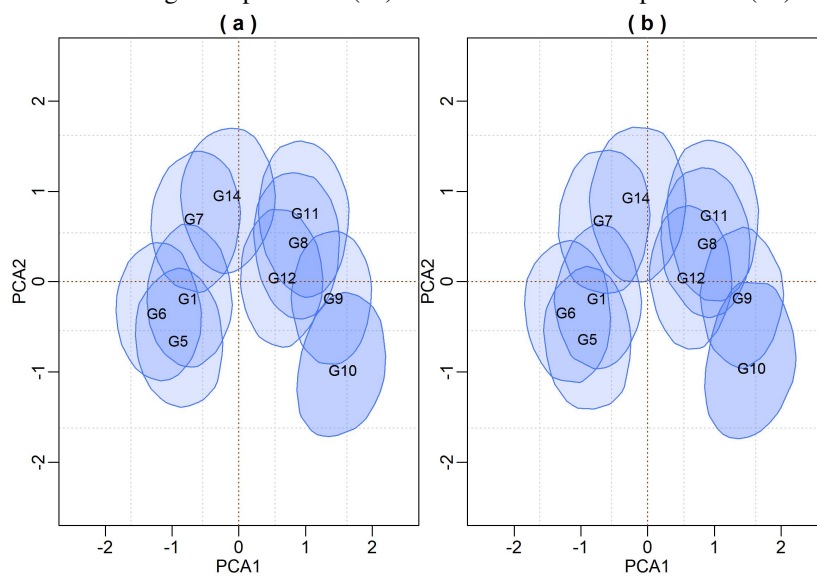


Fonte: Do autor (2019)

4.3.1.3 Modelo BAMMIE-RJ

Analogamente ao realizado nas análises anteriores apresentamos os biplots condicional e marginal para o modelo BAMMIE (Figura 12). Como pode ser visto biplots marginal e condicional também apresentam o mesmo padrão.

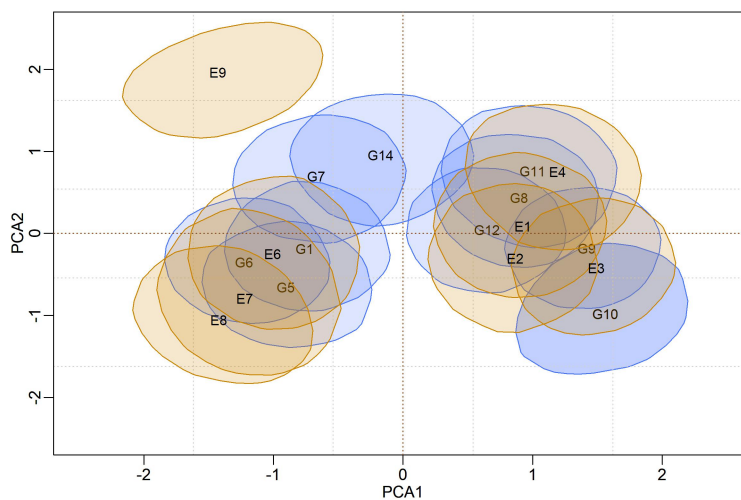
Figura 12 Regiões de credibilidades para o modelo BMMIE, construídas com base na média marginal a posteriori (a) e média condicional a posteriori (b).



Fonte: Do autor (2019)

O biplot completo incluindo regiões de credibilidade para escores genotípicos e ambientais é apresentado pela Figura 13. Interpretações semelhantes às feitas para BAMMI e BMMIE podem ser estabelecidas, observando apenas que os genótipos G2, G3, G4, G13, G15, G16, G17, G18, G19, G20 são considerados estáveis (englobam a origem).

Figura 13 Regiões de credibilidades para os escores genotípicos e ambientais do modelo BAMMIE cujas regiões não contêm a origem (0,0).



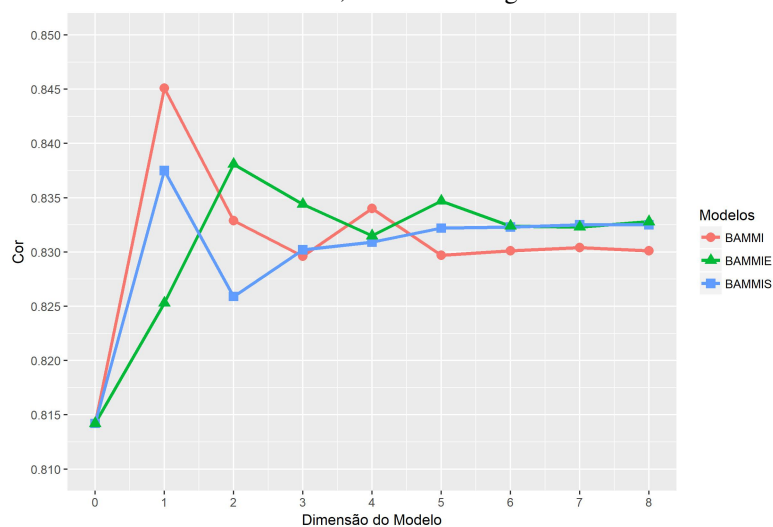
Fonte: Do autor (2019)

4.4 Avaliação da capacidade preditiva dos modelos

Como especificado em métodos a avaliação da capacidade de predição dos modelos foi realizada pelo cálculo da correlação entre valor observado e predito e pela estatística PRESS, sob desbalanceamento aleatório de 10% dos genótipos em ambientes.

Na Figura 14 observa-se o comportamento das três versões do modelo AMMI no que se refere a correlação média entre valores observados e preditos (quanto maior o valor, melhor) para modelos da família AMMI. Os resultados sugerem alta capacidade preditiva do modelo AMMI Bayesiano (BAMMI, BAMMIS e BAMMIE) sob o desbalanceamento sugerido (sempre acima de 81%) com destaque para o BAMMI1 com 83% de correlação.

Figura 14 Avaliação média da capacidade dos modelos com base na Correlação entre o valor observado e o estimado, utilizando o algoritmo Gibbs.

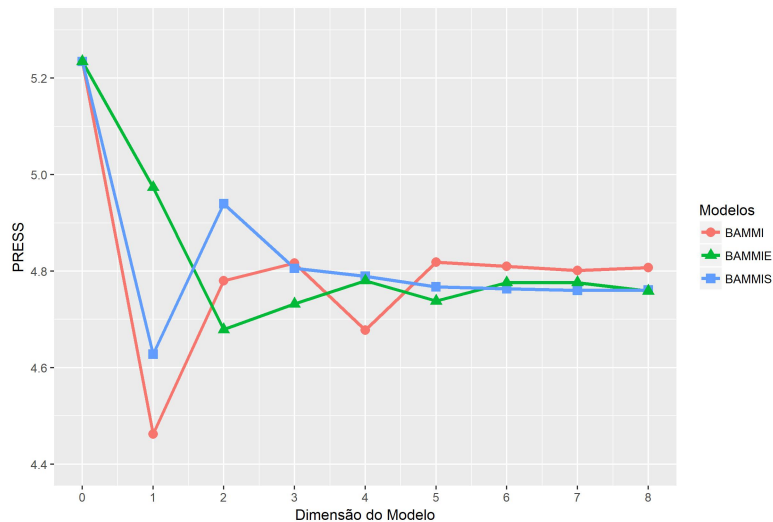


Fonte: Do autor (2019)

Por outro lado, a Figura 15 apresenta graficamente os valores da PRESS, em que menores valores indicam maior capacidade. Por esse critério verificam-se diferenças mais acentuadas entre os modelos compostos apenas por efeitos principais (AMMI0) em relação aos demais. Observa-se também menores valores PRESS para BAMMI1, Bammis1 e Bammie2.

De forma geral, com base nos resultados da análise de validação cruzada (PRESS e COR), os modelos selecionados seriam BAMMI1, Bammis1 e Bammie2, respectivamente. Notando que o modelo BAMMI1 apresentou resultados ligeiramente melhores nas duas medidas utilizadas como critérios de avaliação.

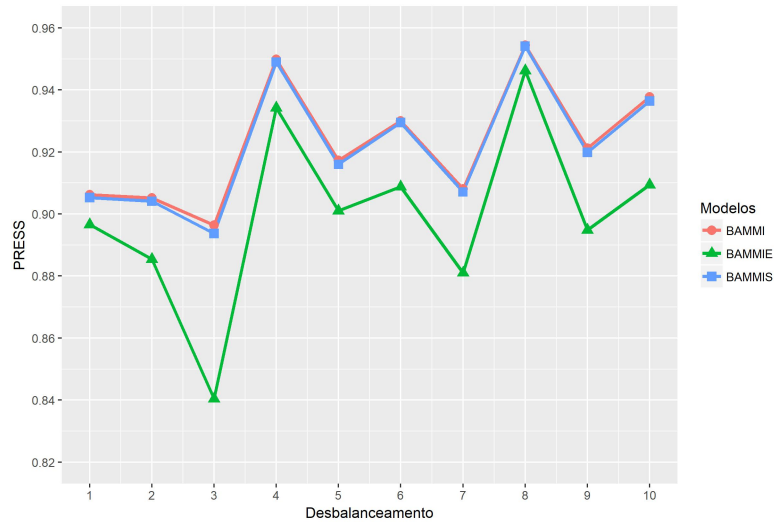
Figura 15 Avaliação média da capacidade dos modelos com base na PRESS, utilizando o algoritmo Gibbs.



Fonte: Do autor (2019)

A avaliação da predição dos modelos utilizando o algoritmo *Reversible Jump*, foi dada com base na dimensão obtida em todos os estados da cadeia, sendo dessa forma a predição ponderada pelo número de eixo retido a cada interação (avaliação marginal). Os resultados da correlação entre valor observado e predito estão descritos na Figura 16 e apontam correlação média sempre acima de 89%. Para *folds* individuais observaram-se valores sempre superiores a 84% e de forma geral observou-se melhor desempenho para BAMMI e BAMMIS.

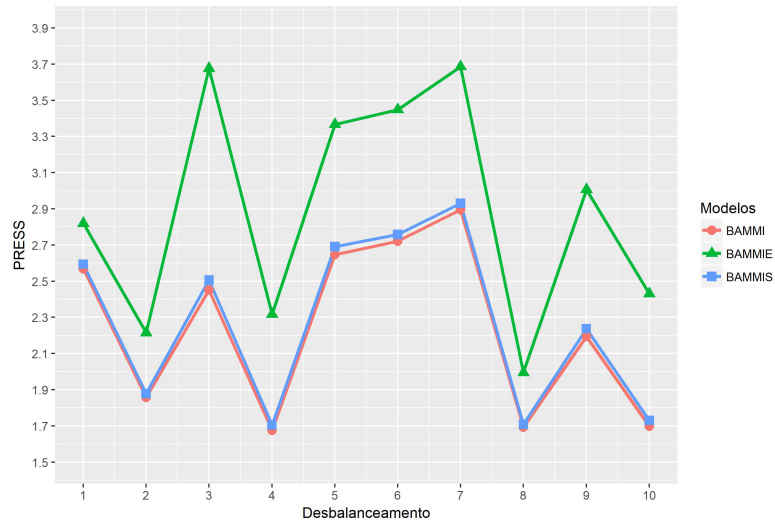
Figura 16 Avaliação da capacidade dos modelos com base na Correlação entre o valor observado e o estimado, utilizando o algoritmo *Reversible Jump*.



Fonte: Do autor (2019)

No que se refere a PRESS (Figura 17) observam-se também desempenhos mais expressivos de BAMMI e Bammis em relação a Bammie, e menores diferenças entre BAMMI e Bammis. Apesar das diferenças não expressivas entre os dois últimos modelos referidos, menores valores foram observados para BAMMI em todos os *folds* individuais. A Tabela com os valores da PRESS e Cor (Tabela 11) está no Apêndice B e oferecem uma melhor discriminação.

Figura 17 Avaliação da capacidade dos modelos com base na PRESS, utilizando o algoritmo Reversible Jump.



Fonte: Do autor (2019)

5 DISCUSSÃO

Muitas pesquisas foram realizadas levando-se em conta as soluções de mínimos quadrados utilizadas nas abordagens AMMI clássicas, em particular o desenvolvimento de procedimentos visando à determinação do número de termos bilineares a serem mantidos no modelo para descrever o efeito da interação GEI. Em abordagens frequentistas, testes baseados na distribuição F e métodos baseados em computação intensiva adquiriram grande aplicabilidade (DIAS; KRZANOWSKI, 2003; GAUCH, 1988; PIEPHO, 1995). A questão de qual a quantidade de termos multiplicativos são necessários para explicar o padrão GEI surge naturalmente na perspectiva bayesiana.

Foi realizada uma análise para dados simulados com o objetivo de utilizar critérios de informação para a seleção de modelos AMMI e, ainda, considerar esse problema na perspectiva do método de saltos reversíveis em cadeias de Markov (GREEN, 1995). O algoritmo *Reversible Jump* é um dos mais populares métodos envolvendo cadeias de Markov transdimensionais e tem sido aplicado em diversos contextos (BALESTRE; SOUZA, 2016; MONDAL et al., 2010; XU, 2013; ZHANG et al., 2004). Em nossa abordagem, como se trata de modelos aninhados, tem-se que o jacobiano da transformação possui valor igual a um (mapeamento identidade). A determinação do mapeamento para casos onde essa condição não é verificada se torna mais complexa e as escolhas dependem das especificidades do problema abordado.

O estudo referente à capacidade preditiva, utilizando validação cruzada, apontou boas propriedades para as três versões AMMI-Bayesiano utilizadas. No que se refere à correlação (Cor) de modelos AMMI bayesianos ajustados com amostrador de Gibbs, verificou-se correlação média mínima de 81%, sendo que os melhores modelos foram BAMMI1, BMMIE2 e BMMIS1 (Figura 14), com

0,8451, 0,8381 e 0,8375, respectivamente, embora observam-se resultados pouco discrepantes para todos os membros da família AMMI.

Por outro lado, os modelos obtidos pelo *Reversible Jump*, apresentaram Cor média (marginal) de 0,9227, 0,9215 e 0,8998, para BAMMI3, BMMIS3 e BMMIE2, respectivamente. Como é possível perceber, os valores da Cor média condicional para os modelos mais prováveis (método *Reversible Jump*) foram sempre maiores para todas as versões AMMI-bayesiano propostas neste estudo. De acordo com a PRESS média (Tabela 11, APÊNDICE B) os modelos ajustados por *Reversible Jump* apresentaram valores expressivamente menores em relação aos modelos ajustados pelo método MCMC convencional (Figura 15).

Determinar o número de termos bilineares a serem mantidos no modelo AMMI sob a ótica bayesiana tem sido pouco abordado na literatura. Liu (2001) oferece uma rica discussão a esse respeito, abordando esse problema por meio de critérios de informação BIC e AIC, e ainda, utilizando o fator de Bayes (BF). O autor supracitado comparou esses métodos entre si e com abordagens frequentistas. Outra aplicação envolvendo a determinação do melhor modelo na análise AMMI é encontrado em Perez-Elizalde, Jarquin e Crossa (2012), que utilizaram o fator de Bayes.

Silva et al. (2015) também abordaram esse tema utilizando estimadores de encolhimento para os valores singulares por meio da escolha de prioris específicas para $\sigma_{\lambda_k}^2$. Essa abordagem pode ser considerada uma versão bayesiana dos métodos propostos por Cornelius e Crossa (1999). O método *shrinkage* de Cornelius e Crossa (1999) dispensa o uso de qualquer critério para seleção de parâmetros bilineares. Entretanto o método envolve a utilização de um preditor cuja expressão se assemelha um teste de análise de variância (teste F) para os valores singulares dado por $S_k = 1 - (u_k \sigma^2 / n E[\hat{\lambda}_k^2])$, embora não conduza a modelos mais parci-

moniosos no que se refere ao número de termos multiplicativos a serem retidos. Nesse sentido, surge a necessidade de se pensar em procedimentos para seleção de modelos na abordagem AMMI com diferentes prioris para variância dos valores singulares (inclusive para AMMIS ou AMMIE) - o que também foi salientado por Silva et al. (2015). Abordagens mais recentes não se preocupam com a seleção de modelos e oferecem avaliações de estabilidade e adaptabilidade por meio de regiões de credibilidades em biplots envolvendo os dois primeiros eixos principais de forma arbitrária (CROSSA et al., 2011; OLIVEIRA et al., 2015, 2016; JÚNIOR et al., 2018)

Apesar da popularidade, critérios como AIC e BIC e, ainda, procedimentos de validação cruzada exigem uma avaliação exaustiva de todos os modelos possíveis, inviabilizando aplicações quando o espaço do modelo é grande, sendo, portanto relevantes apenas em circunstâncias limitadas. Além disso, o fator de Bayes, como argumenta Liu (2001), é sensível à escolha de prioris e, escolher a priori ideal é uma tarefa difícil (embora não se tenha utilizado esse critério na presente abordagem). Nota-se também que diferentes critérios tendem, em geral a escolher modelos diferentes (Figura 1).

Outro ponto a ser destacado é que foram utilizadas diferentes prioris relacionadas ao parâmetro $\sigma_{\lambda_k}^2$. Prioris utilizadas por Crossa et al. (2011) e Oliveira et al. (2015) podem ser consideradas como uma aproximação do princípio da razão insuficiente (Laplace). A versão BAMMIS de Silva et al. (2015), por outro lado, tem como limite a priori de Jeffreys (1961), ou seja, se baseia no princípio da invariância. Além disso, diversas prioris para variância dos valores singulares foram utilizadas por Liu (2001) e podem conduzir a diferentes resultados.

Apesar da ampla utilização de prioris não informativas, as mesmas podem não representar em suma o total de “ignorância” sobre o problema, sendo usadas,

na maioria das vezes, apenas como distribuições de referência, ou seja, uma opção padrão na ausência de qualquer informação disponível. Nesse sentido, existem algumas prioris (ditas não-informativas) que são mais úteis. Contudo, seria abusivo dizer que essas são menos informativas do que outras. Nesse trabalho, foi proposto o uso de prioris baseadas no princípio da máxima entropia (JAYNES, 1957) com o apelo de encontrar a distribuição de probabilidades mais “honesta” possível dadas as informações ou restrições do problema.

No que se refere à análise biplot, apresenta-se o biplot AMMI-2 condicional ao modelo mais provável, bem como com escores marginais. Embora não sendo encontradas discrepâncias expressivas entre as duas versões, optar pelo biplot marginal poderia ser interessante já que produz menor erro de Monte Carlo; por outro lado, o biplot condicional determina maior grau de liberdade para o erro. Além disso, não foram observados grandes diferenças no que se refere ao padrão biplot da interação entre os modelos apresentados. Contudo, quando se considera a análise de estabilidade, a partir das regiões de credibilidade, algumas diferenças quanto a performance das três versões AMMI emergem. A partir das Tabelas 12 e 13 (APÊNDICE B) é possível inferir que os modelos, de modo geral, apresentaram bom desempenho em identificar corretamente genótipos estáveis e não estáveis com base no cenário simulado. Esse comportamento foi verificado tanto para o ajuste bayesiano convencional (amostrador de Gibbs) quanto para o método RJMCMC.

É preciso destacar ainda que a questão da representação biplot dos escores genotípicos e ambientais é um tema controverso. Se critérios determinam o AMMI com um, dois ou três termos bilineares, os biplots em uma, duas ou três dimensões, respectivamente, podem ser utilizados de forma direta. Contudo, para modelos com número de termos mais elevados, a questão se torna mais complexa,

embora existam alguns procedimentos como defendido por Yan e Tinker (2006). Não obstante a essas considerações, a utilização arbitrária do biplot AMMI-2 tem sido utilizada como um procedimento padrão. Apesar do amplo uso dessa ferramenta gráfica, não são encontradas muitas pesquisas estruturais nem discussões mais aprofundadas sobre esses aspectos na literatura.

As questões aqui levantadas representam um bom motivo para optar pelo uso do *Reversible Jump*, embora o conjunto de dados utilizados em nossa abordagem seja pequeno. Nesse sentido, o tempo computacional de análise com *Reversible Jump* é bem mais atraente. Como pode ser observado a partir da Tabela 14, no APÊNDICE B, a seleção de modelos por critérios de informação envolve procedimentos cujo o tempo computacional total (em minutos) são no mínimo quatro vezes maiores em relação ao método RJMCMC. Outra justificativa, como já referido nesse texto, seria a de que esse método leva a uma extensão natural da teoria de cadeias de Markov. De forma geral, considera-se que os resultados aqui apresentados são expressivos e podem contribuir para produzir resultados mais acurados, bem como motivar o uso da análise AMMI bayesiana. Embora essa modelagem já esteja consolidada, contando até com uma biblioteca no software R bayesammi (YASEEN et al., 2018), a aplicação do método ainda é incipiente.

6 CONCLUSÃO

A abordagem AMMI-Bayesiana com o uso do *RJMCMC* permite tanto o ajuste e seleção de modelos, quanto a inferência marginal como média bayesiana de modelos. O método é mais preciso e acurado do que a inferência condicionada à escolha da dimensão (menor PRESS e maior Correlação). Além disso, o *RJMCMC* diminui o tempo computacional para a inferência sobre o número de dimensões. Para a inferência condicional, o AICM é o critério mais consistente, pelo qual se seleciona modelos com mesma dimensão. A escolha de prioris não afeta o modelo marginal da forma que afeta os modelos condicionais. A priori baseada no princípio da máxima entropia (BAMMIE) é escolha parcimoniosa, resultando em modelos de baixa dimensão, mas tende a não alterar as regiões de credibilidade marginais obtidas com as demais prioris. Esta escolha é conveniente no exemplo, pois o melhorista prático tende a preferir resenções biplot.

REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, New York, v. 19, n. 6, p. 716-723, 1974.
- BALESTRE, M.; SOUZA, C. L. de. Bayesian reversible-jump for epistasis analysis in genomic studies. **BMC genomics**, v. 17, n. 1, p. 1012, 2016.
- BALESTRE, M. et al. Yield stability and adaptability of maize hybrids based on gge biplot analysis characteristics. **Crop breeding and applied biotechnology**, Viçosa, v. 9, n. 3, p. 219–228, 2009.
- BERGER, J. O. **Statistical decision theory and Bayesian analysis**. New York: Springer, 1985. 442 p.
- BERNARDO, J. M. Reference posterior distributions for bayesian inference. **Journal of the Royal Statistical Society. Series B Methodological**, London, v. 41, p. 113–147, 1979.
- BERNARDO, J. M.; SMITH, A. F. **Bayesian theory**. New York: J. Wiley, 1994. 597 p.
- BODIN, T.; SAMBRIDGE, M. Seismic tomography with the reversible jump algorithm. **Geophysical Journal International**, v. 178, n. 3, p. 1411–1436, 2009.
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: A. Wesley, 2011. 588 p.
- BURGUEÑO, J. et al. Modeling additive \times environment and additive \times additive \times environment using genetic covariances of relatives of wheat genotypes. **Crop Science**, v. 47, n. 1, p. 311, 2007.
- CARVALHO, C. G. P. d. et al. Interação genótipo \times ambiente no desempenho produtivo da soja no Paraná. **Pesquisa Agropecuária Brasileira**, v. 37, p. 989-1000, 2002.
- CHEN, M.-H.; KUO, L.; LEWIS, P. O. **Bayesian Phylogenetics: methods, algorithms, and applications**. [S.l.]: CRC Press, 2014.
- CORNELIUS, P. et al. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In: _____. **Genotype-by-environment interaction**. Boca Raton: CRC, 1996. p. 199–234.

CORNELIUS, P.; SEYEDSADR, M. Estimation of general linear-bilinear models for two-way tables. **Journal of Statistical Computation and Simulation**, v. 58, n. 4, p. 287–322, 1997.

CORNELIUS, P.; SEYEDSADR, M.; CROSSA, J. Using the shifted multiplicative model to search for “separability” in crop cultivar trials. **Theoretical and Applied Genetics**, v. 84, n. 1/2, p. 161-172, 1992.

CORNELIUS, P. L.; CROSSA, J. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment cultivar trials. **Crop Science**, v. 39, n. 4, p. 998-1009, 1999.

COVER, T. M.; THOMAS, J. A. **Elements of information theory**. [S.l.]: John Wiley & Sons, 2012.

CROSSA, J. Statistical analyses of multilocation trials. **Advances in Agronomy**, v. 44, n. 1, p. 55-85, 1990.

CROSSA, J. et al. Modeling genotype \times environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. **Crop Science**, v. 46, n. 4, p. 1722–1733, 2006.

CROSSA, J. et al. Bayesian estimation of the additive main effects and multiplicative interaction model. **Crop Science**, v. 51, n. 4, p. 1458-1469, 2011.

CROSSA, J.; YANG, R.-C.; CORNELIUS, P. L. Studying crossover genotype \times environment interaction using linear-bilinear models and mixed models. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 9, n. 3, p. 362–380, 2004.

CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. [S.l.]: Universidad Federal de Viçosa, 2006.

DELLAPORTAS, P.; FORSTER, J. J.; NTZOUFRAS, I. On bayesian model and variable selection using mcmc. **Statistics and Computing**, v. 12, n. 1, p. 27–36, 2002.

DIA, M.; WEHNER, T. C.; ARELLANO, C. Analysis of genotype \times environment interaction (g \times e) using SAS programming. **Agronomy Journal**, v. 108, n. 5, p. 1838–1852, 2016.

DIAS, C. T. dos S.; KRZANOWSKI, W. J. Model selection and cross validation in additive main effect and multiplicative interaction models. **Crop Science**, v. 43, n. 3, p. 865-873, 2003.

- DUARTE, J. B.; VENCOVSKY, R. **Interação genótipos x ambientes: uma introdução à análise “AMMI”**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 60 p. (Série Monografias, 9).
- EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties1. **Crop Science**, v. 6, n. 1, p. 36–40, 1966.
- FINLAY, K.; WILKINSON, G. The analysis of adaptation in a plant-breeding programme. **Australian journal of agricultural research**, v. 14, n. 6, p. 742–754, 1963.
- GABRIEL, K. R. The biplot graphic display of matrices with application to principal component analysis. **Biometrika**, v. 58, n. 3, p. 453-467, 1971.
- GABRIEL, K. R. Least squares approximation of matrices by additive and multiplicative models. **Journal of the Royal Statistical Society. Series B Methodological**, v. 40, n. 2, p. 186-196, 1978.
- GABRIEL, K. R. Le biplot-outil d’exploration de données multidimensionnelles. **Journal de la Société française de statistique**, v. 143, n. 3/4, p. 5-55, 2002.
- GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. 2nd ed. New York: Chapman & Hall, 2006. 344 p.
- GAUCH, H. G. Model selection and validation for yield trials with interaction. **Biometrics**, v. 44, n. 3, p. 705-715, 1988.
- GAUCH, H. G. Statistical analysis of yield trials by AMMI and GGE. **Crop Science**, v. 46, n. 4, p. 1488-1500, 2006.
- GAUCH, H. G.; PIEPHO, H.-P.; ANNICCHIARICO, P. Statistical analysis of yield trials by AMMI and GGE: Further considerations. **Crop science**, v. 48, n. 3, p. 866–889, 2008.
- GAUCH, H. G.; ZOBEL, W. Ammi analysis of yield trials. In: _____. **Genotype-by-environment interaction**. Boca Raton: CRC, 1996. p. 85-122.
- GELFAND, A. E.; SMITH, A. F. Sampling-based approaches to calculating marginal densities. **Journal of the American Statistical Association**, v. 85, n. 410, p. 398–409, 1990.
- GELMAN, A. et al. **Bayesian data analysis**. 2nd ed. Boca Raton: CRC, 2003. 665 p.

- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, PAMI-6, n. 6, p. 721–741, 1984.
- GOLLOB, H. F. A statistical model which combines features of factor analytic and analysis of variance techniques. **Psychometrika**, v. 33, n. 1, p. 73-115, 1968.
- GREEN, P. J. Reversible jump markov chain monte carlo computation and bayesian model determination. **Biometrika**, v. 82, n. 4, p. 711–732, 1995.
- GREEN, P. J.; MIRA, A. Delayed rejection in reversible jump metropolis-hastings. **Biometrika**, v. 88, n. 4, p. 1035–1053, 2001.
- HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, v. 31, n. 6, p. 1109–1144, 1983.
- HU, Z.; YANG, R.-C. Improved statistical inference for graphical description and interpretation of genotype \times environment interaction. **Crop Science**, v. 53, n. 6, p. 2400–2410, 2013.
- JAYNES, E. T. Information theory and statistical mechanics. **Physical review**, v. 106, n. 4, p. 620, 1957.
- JAYNES, E. T. **Probability theory: The logic of science**. [S.l.]: Cambridge university press, 2003.
- JEFFREYS, H. **Theory of Probability**. 3rd ed. Oxford: Oxford Classic Texts in the Physical Sciences, 1961. 459 p.
- JÚNIOR, B. et al. AMMI bayesian models to study stability and adaptability in maize. **Agronomy Journal**, v. 110, n. 5, p. 1765–1776, 2018.
- KASS, R. E.; RAFTERY, A. E. Bayes factors. **Journal of the American Statistical Association**, v. 90, n. 430, p. 773–795, 1995.
- KEMPTON, R. The use of biplots in interpreting variety by environment interactions. **The Journal of Agricultural Science**, v. 103, n. 1, p. 123–135, 1984.
- KRZANOWSKI, W. Cross-validation in principal component analysis. **Biometrics**, v. 43, n. 3, p. 575-584, 1987.

- LIU, G. **Bayesian computations for general linear-bilinear models.** 2001. 169 p. Thesis (Doctor of Philosophy) — University of Kentucky, Lexington-Kentucky, 2001.
- LOPES, H. F. A note on reversible jump markov chain monte carlo. **Graduate School of Business, The University of Chicago**, 2006.
- LUNN, D. J.; BEST, N.; WHITTAKER, J. C. Generic reversible jump MCMC using graphical models. **Statistics and Computing**, v. 19, n. 4, p. 395–408, 2008.
- MALOSETTI, M.; RIBAUT, J.-M.; EEUWIJK, F. A. van. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. **Frontiers in Physiology**, v. 4, p. 1-17, 2013.
- MANDEL, J. The partitioning of interaction in analysis of variance. **Journal of Research of the National Bureau of Standards, Series B**, v. 73, p. 309–328, 1969.
- MANDEL, J. A new analysis of variance model for non-additive data. **Technometrics**, v. 13, n. 1, p. 1-18, 1971.
- MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Análise multivariada.** New York: Academic, 1979. 521 p.
- MIGON, H. S. dos; GAMERMAN, D. **Statistical inference: an integrated approach.** New York: Arnold, 1999. 279 p.
- MONDAL, A. et al. Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump markov chain monte carlo methods. **Advances in Water Resources**, v. 33, n. 3, p. 241–256, 2010.
- NUVUNGA, J. et al. Factor analysis using mixed models of multi-environment trials with different levels of unbalancing. **Genet. Mol. Res.**, v. 14, n. 4, p. 14262–14278, 2015.
- Oakey, H. et al. Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. **Theoretical and Applied Genetics**, v. 114, n. 8, p. 1319–1332, 2007.
- OLIVEIRA, L. A. de et al. Credible intervals for scores in the ammi with random effects for genotype. **Crop Science**, v. 55, n. 2, p. 465–476, 2015.
- OLIVEIRA, L. D. et al. Bayesian gge biplot models applied to maize multi-environments trials. **Genet. Mol. Res.**, v. 15, n. 2, p. 1–21, 2016.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Serviço de Educação e Bolsas; Fundação Calouste Gulbenkian, 2003. 446 p.

PERASSA, E. A. **Estudo sobre a aplicação de estatística bayesiana e método de máxima entropia em análise de dados**. Dissertation (Mestrado) — Universidade Estadual de Campinas, Campinas, 2007.

PEREZ-ELIZALDE, S.; JARQUIN, D.; CROSSA, J. A general bayesian estimation method of linear–bilinear models applied to plant breeding trials with genotype \times environment interaction. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 17, n. 1, p. 15-37, 2012.

PIEPHO, H. P. Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials. **Theoretical and Applied Genetics**, v. 90, n. 3/4, p. 438-443, 1995.

PIEPHO, H. P. Analyzing genotype-environment data by mixed models with multiplicative terms. **Biometrics**, v. 90, n. 3/4, p. 761-766, 1997.

PIEPHO, H.-P. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. **Theoretical and Applied Genetics**, v. 97, n. 1-2, p. 195–201, 1998.

PIEPHO, H.-P.; MÖHRING, J. Selection in cultivar trials – Is it ignorable? **Crop Science**, v. 46, n. 1, p. 192–201, 2006.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>.

RAFTERY, A. E.; LEWIS, S. M. [Practical markov chain monte carlo]: comment: one long run with diagnostics: implementation strategies for markov chain Monte Carlo. **Statistical Science**, v. 7, n. 4, p. 493-497, 1992.

RAFTERY, A. E. et al. Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In: _____. **Bayesian Statistics**. [S.l.], 2007. v. 8, p. 1-45.

ROBERTS, G. O.; ROSENTHAL, J. S. Harris recurrence of metropolis-within-gibbs and trans-dimensional markov chains. **The Annals of Applied Probability**, v. 16, n. 4, p. 2123–2139, 2006.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, v. 6, n. 2, p. 461-464, 1978.

SILVA, C. P. da et al. A bayesian shrinkage approach for ammi models. **PloS one**, v. 10, n. 7, p. e0131414, 2015.

SILVA, R. R.; BENIN, G. Análises biplot: conceitos, interpretações e aplicações. **Ciência Rural**, v. 42, n. 8, 2012.

SILVA, W. C. J.; DUARTE, J. B. Métodos estatísticos para estudo de adaptabilidade e estabilidade fenotípica em soja. **Pesquisa Agropecuária Brasileira**, v. 41, n. 1, p. 23–30, 2006.

SISSON, S. A. Transdimensional markov chains. **Journal of the American Statistical Association**, v. 100, n. 471, p. 1077–1089, 2005.

SMITH, A.; CULLIS, B.; GILMOUR, A. Applications: the analysis of crop variety evaluation data in Australia. **Australian & New Zealand Journal of Statistics**, v. 43, n. 2, p. 129-145, 2001.

SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC methods in quantitative genetics**. [S.l.]: Springer Science & Business Media, 2007.

SPIEGELHALTER, D. J.; ABRAMS, K. R.; MYLES, J. P. **Bayesian approaches to clinical trials and health-care evaluation**. Chichester: J. Wiley, 2004. 391 p.

TIERNEY, L. Markov chains for exploring posterior distributions. **The Annals of Statistics**, JSTOR, p. 1701–1728, 1994.

TIERNEY, L. A note on metropolis-hastings kernels for general state spaces. **Annals of applied probability**, JSTOR, p. 1–9, 1998.

VIELE, K.; SRINIVASAN, C. Parsimonious estimation of multiplicative interaction in analysis of variance using kullback-leibler information. **Journal of Statistical Planning and Inference**, v. 84, n. 1, p. 201-219, 2000.

WAAGEPETERSEN, R.; SORENSEN, D. A tutorial on reversible jump MCMC with a view toward applications in QTL-mapping. **International Statistical Review**, Wiley, v. 69, n. 1, p. 49–61, 2001.

XU, S. **Principles of statistical genomics**. [S.l.]: Springer, 2013.

YAN, W.; GLOVER, K. D.; KANG, M. S. Comment on “biplot analysis of genotype \times environment interaction: proceed with caution”. **Crop Science**, v. 50, n. 4, p. 1121-1123, 2010.

YAN, W. et al. Cultivar evaluation and mega-environment investigation based on the gge biplot. **Crop Science**, v. 40, n. 3, p. 597–605, 2000.

YAN, W.; KANG, M. S. **GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists**. [S.l.]: CRC press, 2002.

YAN, W.; TINKER, N. A. Biplot analysis of multi-environment trial data: Principles and applications. **Canadian Journal of Plant Science**, v. 86, n. 3, p. 623-645, 2006.

YASEEN, M. et al. **bayesammi: Bayesian Estimation of the Additive Main Effects and Multiplicative Interaction Model**. [S.l.], 2018. R package version 0.1.0. Disponível em: <<https://CRAN.R-project.org/package=bayesammi>>.

ZHANG, Z. et al. Bayesian inference on principal component analysis using reversible jump markov chain monte carlo. In: **AAAI**. [S.l.: s.n.], 2004. p. 372-377.

ZOBEL, R. W.; WRIGHT, M. J.; GAUCH, H. G. Statistical analysis of a yield trial. **Agronomy Journal**, v. 80, n. 3, p. 388-393, 1988.

APÊNDICE A

A seguir é apresentada uma justificativa para uso de priori de entropia para as componentes de variância dos valores singulares.

Priori de máxima entropia

Uma das etapas fundamentais do método bayesiano é a atribuição de prioris para quantidades desconhecidas, em especial aos parâmetros, que podem ser declaradas por meio de densidades de probabilidade. O princípio da razão insuficiente de Laplace e prioris de Jeffreys adquiriram grande aplicabilidade quando se deseja considerar “ignorância a priori” para algum parâmetro e são referidas como prioris objetivas, pela abordagem se basear fundamentalmente na função de verossimilhança. Por outro lado, prioris mais informativas podem ser atribuídas de forma subjetiva, utilizando, por exemplo, o critério de famílias conjugadas, amparadas pelas facilidades algébricas, já que priori e posteriori pertencem à mesma família paramétrica e a atualização da informação envolve apenas uma mudança nos valores de hiperparâmetros. A questão é que diferentes escolhas podem levar a resultados expressivamente diferentes para uma mesma verossimilhança.

Diferentes prioris foram atribuídas para os componentes de variância $\sigma_{\lambda_k}^2$ do modelo AMMI-Bayesiano (Liu, 2001, Crossa et al., 2011, Silva et al. 2015), resultando em diferenças expressivas nas estimativas dos vetores singulares, bem como em termos inferenciais. Com base nessas informações, foi utilizado aqui o princípio da máxima entropia para incorporar priori $\sigma_{\lambda_k}^2$. O princípio da entropia máxima, enquanto método de inferência estatística (JAYNES, 1957), leva à seleção de uma função de densidade de probabilidade consistente com conhecimento que se tem e capaz de produzir estimativas menos tendenciosas o possível sobre a informação fornecida.

A medida de entropia para uma variável contínua y é quantificada por uma função $H(\cdot)$ dada por:

$$H(y) = - \int_{-\infty}^{\infty} f(y) \log(f(y)) dy.$$

Para encontrar a distribuição de máxima entropia deve-se maximizar $H(y)$ sob todas as distribuições $f(y)$ que satisfazem:

1. $f(y) \geq 0$, com a igualdade válida somente fora do domínio S da variável;
2. $\int_S f(y) dy = 1$;
3. $\int_S f(y) g_i(y) dy = a_i$, para $1 \leq i \leq m$; f é uma no conjunto suporte S que atende a determinadas restrições de momentos.

Para resolver o problema acima, é necessário utilizar os multiplicadores de Lagrange. Assim, pode-se escrever o seguinte Lagrangiano (COVER; THOMAS, 2012):

$$J(f(y)) = - \int_S f(y) \ln(f(y)) dy + \beta_0 \left(\int_S f(y) dy \right) + \quad (32)$$

$$+ \sum_{i=1}^m \beta_i \left(\int_S f(y) g_i(y) dy = a_i \right)$$

em que β_0, \dots, β_i são os multiplicadores de Lagrange.

Derivando a equação (32) em relação à distribuição $f(y)$ tem-se:

$$\frac{\partial J(f(y))}{\partial f(y)} = -\ln(f(y)) - 1 + \beta_0 + \sum_{i=1}^m \beta_i g_i(y)$$

em que β_0, \dots, β_i são escolhidos de tal forma que $f(y)$ satisfaça as restrições.

Priori de máxima entropia para $\sigma_{\lambda_k}^2$

Nosso objetivo aqui é construir uma priori baseada no principio máxima entropia para a componente de variância de cada $\sigma_{\lambda_k}^2$ de cada λ_k . A única restrição imposta admitida é que $E(\sigma_k^2) \geq 0$.

Para facilitar cálculos algébricos optou-se por trabalhar com a precisão $\tau_{\lambda_k} = \frac{1}{\sigma_{\lambda_k}^2}$. Além disso, tem-se que: $E(\sigma_{\lambda_k}^2) \geq 0 \rightarrow E(\tau_{\lambda_k}) \geq 0$.

Utilizando a função lagrangiana temos que a função que maximiza a entropia é:

$$\int_0^{+\infty} \tau_{\lambda_k} p(\tau_{\lambda_k}) d\tau_{\lambda_k} = \vartheta \rightarrow \Lambda \left[\int_0^{+\infty} \tau_{\lambda_k} p(\tau_{\lambda_k}) - \vartheta \right].$$

A expressão geral da distribuição de máxima entropia gerada a partir do vínculo considerado é (PERASSA, 2007):

$$P(\tau_{\lambda_k}) = \frac{e^{-\Lambda\tau_{\lambda_k}}}{\int_0^{+\infty} e^{-\Lambda\tau_{\lambda_k}} d\tau_{\lambda_k}} \quad (33)$$

Observe que:

$$\int_0^{+\infty} e^{-\Lambda\tau_{\lambda_k}} d\tau_{\lambda_k} = \int_0^{+\infty} \frac{\Lambda}{\Lambda} e^{-\Lambda\tau_{\lambda_k}} d\tau_{\lambda_k} = \frac{1}{\Lambda} \int_0^{+\infty} \Lambda e^{-\Lambda\tau_{\lambda_k}} d\tau_{\lambda_k} = \frac{1}{\Lambda}; \Lambda > 0.$$

Assim:

$$P(\tau_{\lambda_k}) = \frac{e^{-\Lambda\tau_{\lambda_k}}}{1/\Lambda} = \Lambda e^{-\Lambda\tau_{\lambda_k}}.$$

Então encontramos a densidade da variável τ_{λ_k} . Contudo, estamos interessados na densidade de $\sigma_{\lambda_k}^2$ e $\sigma_{\lambda_k}^2 = 1/\tau_{\lambda_k}$. A densidade de $\sigma_{\lambda_k}^2$ pode ser obtida pelo método do jacobiano

$$f_{\sigma_{\lambda_k}^2}(\sigma_{\lambda_k}^2) = \left| \frac{d}{d\sigma_{\lambda_k}^2} g^{-1}(\sigma_{\lambda_k}^2) \right| f_{\tau_{\lambda_k}}(g^{-1}(\sigma_{\lambda_k}^2)) I_D(\sigma_{\lambda_k}^2)$$

com: $\sigma_{\lambda_k}^2 = g(\tau_{\lambda_k}) = 1/\tau_{\lambda_k} \Rightarrow (\sigma_{\lambda_k}^2)^{-1} = \tau_{\lambda_k}$,

então:

$$\frac{d}{d\sigma_{\lambda_k}^2} g^{-1}(\sigma_{\lambda_k}^2) = \frac{d}{d\sigma_{\lambda_k}^2} \left(\frac{1}{\sigma_{\lambda_k}^2} \right) = -1/(\sigma_{\lambda_k}^2)^2.$$

Logo:

$$f(\sigma_{\lambda_k}^2) = \Lambda e^{-\Lambda/\sigma_{\lambda_k}^2} \left| -1/(\sigma_{\lambda_k}^2)^2 \right| = \frac{\Lambda}{\Gamma(1)} (\sigma_{\lambda_k}^2)^{-(1+1)} e^{-\Lambda/\sigma_{\lambda_k}^2};$$

sendo: $\Gamma(\cdot)$ é a função gama.

Assim: $\sigma_{\lambda_k}^2 \sim Inv - gama(a, b)$ em que $a = 1$ e $b = \Lambda$ e $\Lambda > 0$.

APÊNDICE B

A seguir é apresentada algumas tabelas, para melhor descrever alguns dos resultados já apresentados na seção Resultados.

Tabela 5 Estimativas pontuais para a componente de variância do erro e da variância do efeito genotípico, referente ao modelo BAMMI, para as possíveis dimensões, utilizando algoritmo Gibbs.

Mod.	(σ_e^2)				(σ_g^2)			
	Média	Dp	LI	LS	Média	Dp.	LI	LS
1	3,433	0,223	5,055	20,530	11,077	4,437	5,055	20,529
2	2,958%	0,196	2,589	3,357	11,148	4,444	5,320	20,733
3	2,753&	0,197	2,403	3,149	11,077	4,337	5,638	20,724
4	2,689*	0,201	2,318	3,105	11,042	4,308	5,702	20,956
5	2,711	0,210	2,342	3,159	11,000	4,428	5,053	20,383
6	2,722	0,209	2,346	3,149	11,179	4,487	5,606	21,289
7	2,737	0,211	2,337	3,152	10,977	4,369	5,218	20,451
8	2,733	0,209	2,338	3,155	11,084	4,507	5,269	20,828

Os modelos vencedores com base nos critérios de informações são designados por: “ * ” AIC; “ & ” BIC e “ % ” AICM.

Tabela 6 Estimativas pontuais para a componente de variância do erro e da variância do efeito genotípico, referente ao modelo BAMMIS, para as possíveis dimensões, utilizando algoritmo Gibbs.

Mod.	(σ_e^2)				(σ_g^2)			
	Média	Dp.	LI	LS	Média	Dp.	LI	LS
1	3,435	0,225	2,983	3,871	11,146	4,333	5,499	20,783
2	2,96%	0,204	2,587	3,369	11,238	4,395	5,266	20,737
3	2,76*&	0,195	2,411	3,171	11,154	4,315	5,190	20,540
4	2,775	0,238	2,346	3,252	11,176	4,55	5,206	20,675
5	2,865	0,221	2,442	3,302	11,122	4,255	5,539	20,587
6	2,864	0,221	2,441	3,293	11,088	4,289	5,140	20,160
7	2,869	0,226	2,458	3,328	11,091	4,406	5,383	20,439
8	2,870	0,224	2,459	3,325	11,104	4,398	5,242	20,603

Os modelos vencedores com base nos critérios de informações são designados por: “ * ” AIC;
“ & ” BIC e “ % ” AICM.

Tabela 7 Estimativas pontuais para a componente de variância do erro e da variância do efeito genotípico, referente ao modelo BAM-MIE, para as possíveis dimensões, utilizando algoritmo Gibbs.

Mod.	(σ_e^2)				(σ_g^2)			
	Média	Dp.	LI	LS	Média	Dp.	LI	LS
1	3,438	0,227	3,001	3,896	11,079	4,527	5,337	20,557
2	2,965*%&	0,201	2,581	3,366	11,118	4,539	5,317	21,446
3	3,428	0,238	2,963	3,913	10,962	4,321	5,257	20,244
4	3,428	0,239	2,945	3,899	11,038	4,455	5,129	20,495
5	3,433	0,239	2,974	3,922	11,125	4,482	5,296	20,577
6	3,418	0,249	2,932	3,933	11,045	4,536	5,265	21,006
7	3,427	0,242	2,976	3,924	11,075	4,465	5,326	20,717
8	3,447	0,228	3,003	3,897	12,003	4,476	5,260	20,452

Os modelos vencedores com base nos critérios de informações são designados por: “ * ” AIC;
“ & ” BIC e “ % ” AICM.

Tabela 8 Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional (BAMMI3) e marginal para o modelo BAMMI.

Gen.	Condicional			Marginal		
	Média	LI	LS	Média	LI	LS
G1	2,3076	0,6214	3,9259	2,3041	0,6162	3,9388
G2	-4,2906	-6,0000	-2,6143	-4,2994	-5,9613	-2,6143
G3	-0,7547	-2,3944	0,9580	-0,7551	-2,4082	0,9487
G4	3,8307	2,0673	5,4454	3,8230	2,1555	5,5205
G5	-2,6039	-4,3413	-0,9965	-2,6074	-4,3154	-0,9471
G6	-1,1793	-2,9050	0,4380	-1,1813	-2,8959	0,4271
G7	0,6503	-0,9661	2,3024	0,6504	-0,9592	2,3860
G8	0,5387	-1,1076	2,3032	0,5334	-1,0954	2,2990
G9	-3,5352	-5,2939	-1,9268	-3,5394	-5,3327	-1,9399

G10	3,9510	2,3028	5,6735	3,9479	2,3028	5,6836
G11	5,9657	4,3418	7,6942	5,9612	4,1554	7,5331
G12	3,7171	2,0483	5,3837	3,7208	2,0365	5,4009
G13	4,3909	2,7733	6,0759	4,3869	2,7277	6,0655
G14	-2,0399	-3,6915	-0,3521	-2,0432	-3,6216	-0,2895
G15	1,4976	-0,1610	3,1490	1,4970	-0,1531	3,1748
G16	-3,6424	-5,2962	-1,9753	-3,6448	-5,3188	-1,9753
G17	0,3446	-1,3072	2,0240	0,3447	-1,3111	2,0420
G18	-2,8979	-4,6051	-1,2394	-2,9004	-4,6086	-1,2490
G19	-5,2065	-6,8231	-3,4170	-5,2125	-6,9465	-3,5473
G20	-1,2811	-3,0541	0,3816	-1,2832	-2,9417	0,4753

Tabela 9 Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional (BAMMIS3) e marginal para o modelo BAMMIS.

Gen.	Condicional			Marginal		
	Média	LI	LS	Média	LI	LS
G1	2,3114	0,6842	3,9948	2,3143	0,7138	4,0356
G2	-4,2969	-5,8495	-2,6065	-4,2863	-5,8741	-2,6254
G3	-0,7452	-2,3312	0,9947	-0,7382	-2,3473	0,9756
G4	3,8311	2,1339	5,4486	3,8406	2,2291	5,5549
G5	-2,6103	-4,2141	-0,9774	-2,6027	-4,2206	-0,9774
G6	-1,1753	-2,7127	0,6196	-1,1643	-2,7602	0,5376
G7	0,6568	-0,9981	2,3626	0,6700	-1,0290	2,3249
G8	0,5414	-1,1303	2,2089	0,5560	-1,1279	2,2095
G9	-3,5288	-5,1106	-1,8403	-3,5228	-5,2240	-1,9176
G10	3,9629	2,2467	5,5759	3,9663	2,3013	5,6334
G11	5,9743	4,3784	7,7055	5,9833	4,3462	7,6585
G12	3,7314	2,0599	5,3234	3,7370	2,0599	5,3331
G13	4,4067	2,8009	6,0859	4,4056	2,7956	6,0949
G14	-2,0387	-3,7136	-0,4307	-2,0261	-3,6334	-0,3477
G15	1,5044	-0,0964	3,2216	1,5180	-0,0995	3,1830
G16	-3,6410	-5,2279	-1,9206	-3,6339	-5,3038	-1,9681
G17	0,3517	-1,3504	1,9632	0,3559	-1,3740	1,9588
G18	-2,8874	-4,5102	-1,2513	-2,8822	-4,5265	-1,2571
G19	-5,2048	-6,8430	-3,4962	-5,1996	-6,7997	-3,4706

G20	-1,2709	-2,8956	0,3985	-1,2633	-2,8877	0,4205
-----	---------	---------	--------	---------	---------	--------

Tabela 10 Médias a posteriori e regiões a 95% de credibilidade para efeito de genótipos, para o modelo condicional (BAMMIE2) e marginal para o modelo BAMMIE.

Gen.	Condicional			Marginal		
	Média	LI	LS	Média	LI	LS
G1	2,2943	0,5728	3,9539	2,2958	0,5771	3,9343
G2	-4,3083	-6,0147	-2,5553	-4,3048	-6,0290	-2,5980
G3	-0,7602	-2,4000	1,0279	-0,7637	-2,4741	0,9234
G4	3,8139	2,0032	5,4546	3,8127	2,1224	5,5062
G5	-2,6277	-4,2318	-0,8227	-2,6258	-4,3294	-0,9568
G6	-1,1901	-2,8991	0,5450	-1,1844	-2,8898	0,5199
G7	0,6466	-1,0278	2,4284	0,6466	-1,1320	2,2881
G8	0,5237	-1,2759	2,2116	0,5240	-1,2600	2,2124
G9	-3,5504	-5,2802	-1,8556	-3,5464	-5,3332	-1,9530
G10	3,9432	2,2326	5,6365	3,9410	2,2326	5,6067
G11	5,9487	4,1586	7,6092	5,9525	4,1979	7,5909
G12	3,7083	1,9984	5,4686	3,7062	1,9833	5,3911
G13	4,3856	2,7999	6,2159	4,3845	2,6294	6,0117
G14	-2,0419	-3,7523	-0,2517	-2,0422	-3,7274	-0,2724
G15	1,4886	-0,2357	3,2118	1,4915	-0,2737	3,1121
G16	-3,6523	-5,3489	-1,8894	-3,6533	-5,3489	-1,9472
G17	0,3288	-1,3952	2,0459	0,3339	-1,3952	2,0186
G18	-2,9104	-4,6059	-1,2256	-2,9125	-4,6125	-1,2469
G19	-5,2230	-7,1012	-3,6650	-5,2212	-6,8847	-3,4943
G20	-1,2826	-2,9213	0,5048	-1,2842	-2,8635	0,5048

Tabela 11 Correlação e Press para as diferentes suposições feitas sob o modelo AMMI (Reversible Jump).

Desb.	BAMMI		BAMMIS		BAMMIE	
	Cor	Press	Cor	Press	Cor	Press

1	0,9062	2,5688	0,9053	2,5941	0,8966	2,8199
2	0,9052	1,8575	0,9041	1,8785	0,8854	2,2149
3	0,8964	2,4518	0,8937	2,5061	0,8404	3,6766
4	0,9499	1,6744	0,9490	1,7055	0,9342	2,3175
5	0,9173	2,6461	0,9160	2,6913	0,9010	3,3667
6	0,9300	2,7205	0,9295	2,7585	0,9088	3,4480
7	0,9081	2,8939	0,9071	2,9307	0,8810	3,6857
8	0,9544	1,6925	0,9541	1,7080	0,9462	1,9956
9	0,9212	2,1946	0,9199	2,2376	0,8948	3,0070
10	0,9378	1,6982	0,9364	1,7305	0,9094	2,4309
Média	0,9227	2,2398	0,9215	2,2741	0,8998	2,8963

Tabela 12 Classificação do padrão dos escores genotípicos pela representação biplot AMMI2 com base no cenário simulado (algoritmo Gibbs padrão).

Instavel										
Modelo	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
BAMMI	V	F	V	F	V	V	V	V	V	V
BAMMIE	V	F	F	F	V	V	V	V	V	V
BAMMIS	V	F	V	F	V	V	V	V	V	V
Estavel										
Modelo	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20
BAMMI	F	F	F	F	V	F	V	V	V	V
BAMMIE	F	V	F	F	V	F	V	V	V	V
BAMMIS	F	F	V	F	V	V	V	V	V	V

Tabela 13 Classificação do padrão dos escores genotípicos pela representação biplot AMMI2 com base no cenário simulado (algoritmo RJMCMC).

Instavel										
Modelo	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
BAMMI	V	F	V	F	V	V	V	V	V	V
BAMMIE	V	F	F	F	V	V	F	V	V	V
BAMMIS	V	F	F	F	V	V	V	V	V	V

Estavel										
Modelo	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20
BAMMI	F	F	F	F	V	F	V	V	V	V
BAMMIE	F	F	V	F	V	V	V	V	V	V
BAMMIS	F	F	V	F	V	V	V	V	V	V

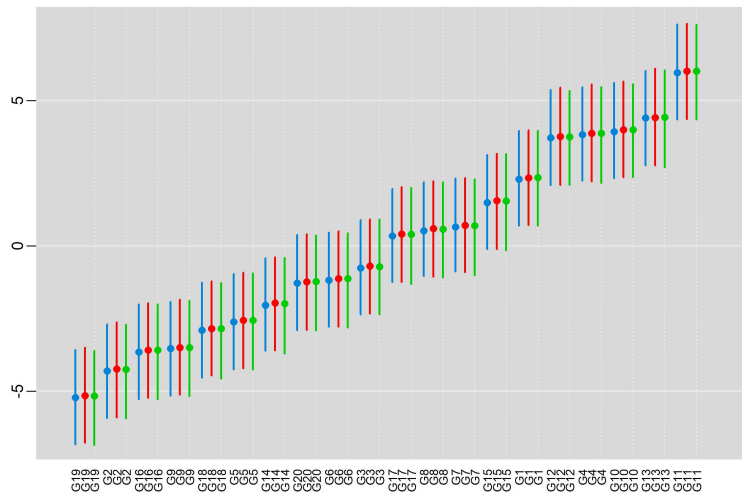
Tabela 14 Tempo computacional total, em minutos (mins), para o processo de amostragem das três versões AMMI-Bayesiano considerando ajuste por critérios de informação e pelo método RJMCMC.

Modelo	Tempo	
	MCMC / Critérios de informação	RJMCMC
BAMMI	887,6 mins	210,19 mins
BAMMIE	913,87 mins	190,24 mins
BAMMIS	905,68 mins	198,84 mins

APÊNDICE C

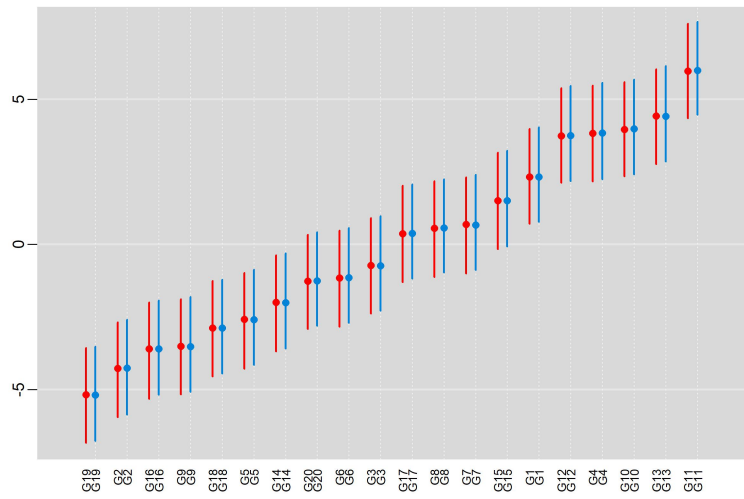
A seguir são apresentadas algumas figuras, para ilustrar alguns resultados não apresentados na seção Resultados.

Figura 18 Efeito genotípico para os modelos Bammie2 (AICM - Vermelho), Bammis3 (BIC - Azul) e BammI4 (AIC - Verde), ajustados com algoritmo de Gibbs.



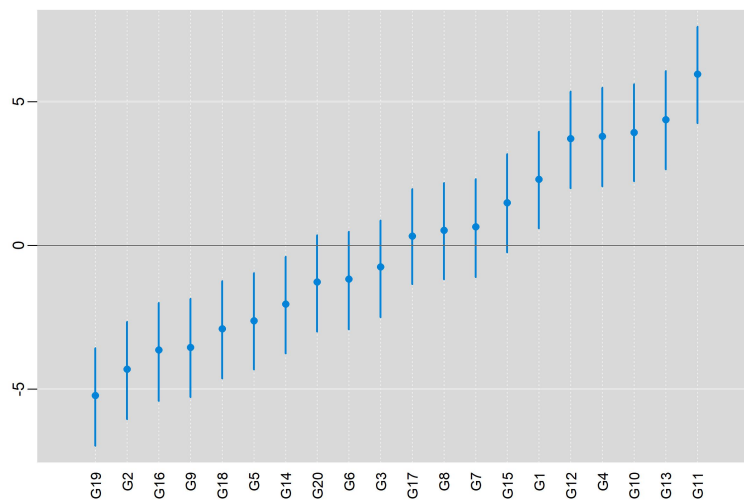
Fonte: Do autor (2019)

Figura 19 Efeito genotípico para os modelos BAMMIS2 (AICM - Vermelho) e BAMMIS3 (AIC/BIC - Azul), ajustados com algoritmo de Gibbs.



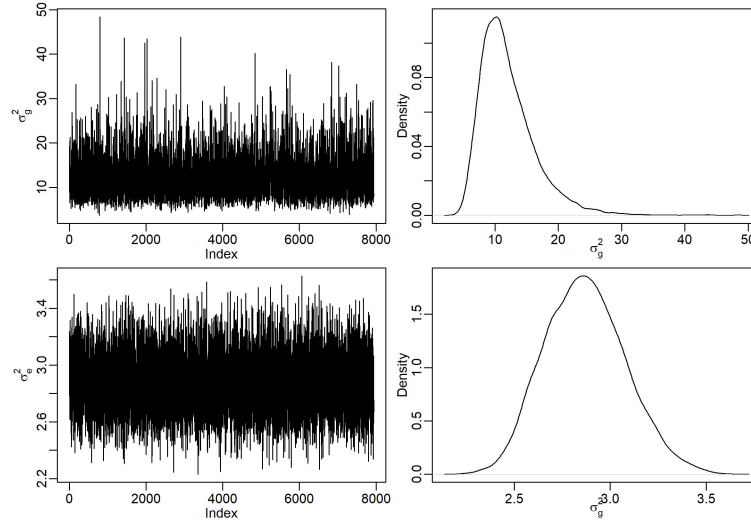
Fonte: Do autor (2019)

Figura 20 Efeito genotípico para o modelo BAMMI2 com base no critério BIC/AIC/AICM, ajustados com algoritmo de Gibbs.



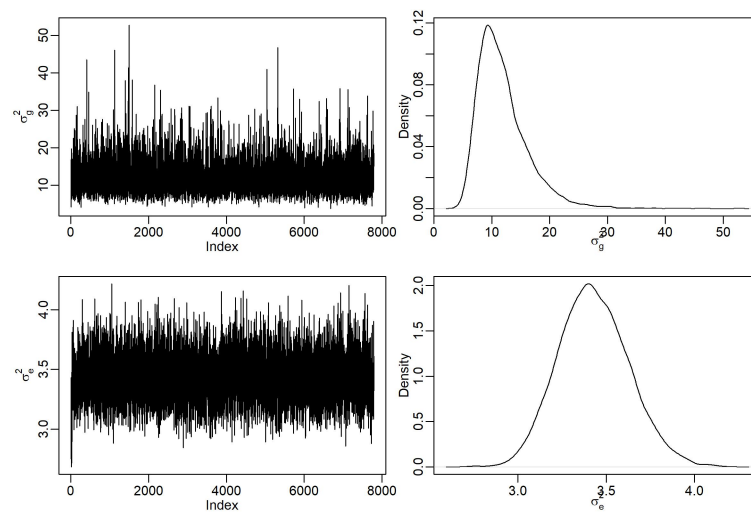
Fonte: Do autor (2019)

Figura 21 Traço e densidades das cadeias para os comentários do modelo BMMIS-RJ.



Fonte: Do autor (2019)

Figura 22 Traço e densidades das cadeias para os comentários do modelo BMMIE-RJ.



Fonte: Do autor (2019)