

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

**ABORDAGEM BAYESIANA E MODELOS MISTOS PARA
EXPERIMENTOS MULTIAMBIENTES NA CULTURA DA
SOJA**

Alysson Jalles da Silva

Engenheiro Agrônomo

2017

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

**ABORDAGEM BAYESIANA E MODELOS MISTOS PARA
EXPERIMENTOS MULTIAMBIENTES NA CULTURA DA
SOJA**

Alysson Jalles da Silva

Orientador: Prof. Dr. Antonio Orlando Di Mauro

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Agronomia (Genética e Melhoramento de Plantas).

2017

S586a Silva, Alysson Jalles da
Abordagem Bayesiana e modelos mistos para experimentos
multiambientes na cultura da soja / Alysson Jalles da Silva. –
Jaboticabal, 2017
iii, 33 p. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de
Ciências Agrárias e Veterinárias, 2017
Orientador: Antonio Orlando Di Mauro
Banca examinadora: Flávio Cese Arantes, João Antonio da Costa
Andrade, Dilermando Perecin, Antonio Luís de Oliveira

Bibliografia

1. *Glycine max*. 2. Melhoramento genético. 3. Modelagem
matemática. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e
Veterinárias.

CDU 633.34:631.52

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação –
Diretoria Técnica de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

CERTIFICADO DE APROVAÇÃO

TÍTULO: ABORDAGEM BAYESIANA E MODELOS MISTOS PARA EXPERIMENTOS
MULTIAMBIENTAIS NA CULTURA DA SOJA

AUTOR: ALYSSON JALLES DA SILVA

ORIENTADOR: ANTONIO ORLANDO DI MAURO

Aprovado como parte das exigências para obtenção do Título de Doutor em AGRONOMIA
(GENÉTICA E MELHORAMENTO DE PLANTAS), pela Comissão Examinadora:

Prof. Dr. ANTONIO ORLANDO DI MAURO
Departamento de Produção Vegetal / FCAV / UNESP - Jaboticabal

Pesquisador Dr. FLÁVIO CESE ARANTES
BAYER S.A. / Londrina/PR

Prof. Dr. JOAO ANTONIO DA COSTA ANDRADE
Departamento de Biologia e Zootecnia / Faculdade de Engenharia de Ilha Solteira

Prof. Dr. DILERMANDO PERECIN
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Prof. Dr. ANTONIO LUÍS DE OLIVEIRA
FAFRAM - Ituverava, SP

Jaboticabal, 06 de abril de 2017.

DADOS CURRICULARES DO AUTOR

Alysson Jalles da Silva – Nascido em 24 de março de 1990 em Teotônio Vilela-Alagoas, filho de Amaro José da Silva (*in memorian*) e Maria Aureliano da Silva. Iniciou seus estudos em ciências agrárias no ano de 2001 na Escola Agrícola São Francisco de Assis (Ensino fundamental) (EASFA, Junqueiro-AL), realizou Ensino médio e curso Técnico em Agropecuária no período de 2005 a 2007 na Escola Agrotécnica Federal de Satuba (EAFS) (Atual: Instituto Federal de Alagoas, Satuba-AL). Fez agronomia no período de fevereiro/2008 a fevereiro/2012, na Universidade Federal de Alagoas, Maceió (UFAL). Durante o bacharelado foi monitor de duas disciplinas: biotecnologia (Docente responsável: Prof. Dr. Cicero Eduardo Ramalho Neto) e forragicultura (Docente responsável: Prof. Dr. José Teodorico de Araújo Filho). Também foi colaborador em 6 projetos do Programa Institucional de Bolsas de Iniciação Científica (PIBIC) na área de melhoramento genético de plantas sob orientação do Prof. Dr. Paulo Vanderlei Ferreira. Além disso, participou de um projeto de extensão desenvolvido no Litoral Norte de Alagoas como bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CONRAD) (Docente responsável: Prof. Dr. José Roberto Santos). Fez o curso de mestrado no período de fevereiro/2012 a julho/2013, na Universidade Federal Rural de Pernambuco, Recife-PE (UFRPE) onde foi bolsista pela Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE), trabalhando com melhoramento de hortaliças e desenvolvendo o projeto “seleção de linhagens de tomate resistentes à *Meloidogyne enterolobii*” sob orientação do Prof. Dr. José Luiz Sandes de Carvalho Filho. Ingressou em agosto/2013 no doutorado em Agronomia (Genética e Melhoramento de Plantas), na Faculdade de Ciências Agrárias e Veterinárias, da Universidade Estadual Paulista “Júlio de Mesquita Filho” – Unesp, Câmpus de Jaboticabal e bolsista CAPES sob orientação do Prof. Dr. Antonio Orlando Di Mauro. Ainda no doutorado o Pós-Graduando teve a oportunidade de trabalhar com diversos tipos de análises estatísticas usando linguagens de programação. Em março de 2017 foi contratado como Analista de dados pela empresa NovAmérica Agrícola Ltda. Em abril de 2017 tornou-se Doutor em Agronomia (Genética e Melhoramento de Plantas) com a tese: “Abordagem Bayesiana e modelos mistos para experimentos multiambientais na cultura da soja”.

Quero conhecer os pensamentos de Deus... O resto é detalhe.

Albert Einstein

Aos meus pais, Amaro José da Silva (*In memoriam*) e Maria Aureliana da Silva que sempre me incentivaram a estudar independentemente da situação. Ao Manoel Alves Cabral pela grande ajuda e por acreditar no meu potencial.

DEDICO

Aos meus irmãos Ariana Maria da Silva e Almir Rudson da Silva, a toda minha família. A minha esposa Janiele Silva da Divincula por estar sempre ao meu lado me fazendo companhia e me apoiando. A minha filhinha de um aninho, Júlia Louise Silva da Divincula a quem amo bastante.

OFEREÇO

AGRADECIMENTOS

A Deus por me permitir alcançar um título e conhecimento que será transmitido para outras gerações, usado para o progresso empresarial, da ciência e desenvolvimento regional.

Ao meu pai Amaro José da Silva (*In memoriam*) pelos conselhos proferidos e pelo exemplo de integridade como pessoa.

A minha mãe Maria Aureliana da Silva por ser minha conselheira, incentivadora, apoiadora e mantenedora contínua.

Ao Manoel Alves Cabral pelo apoio paternal e financeiro em todos os momentos da minha vida acadêmica.

A minha família materna e paterna por sempre acreditar que eu poderia “ser alguém na vida”.

A Faculdade de Ciências Agrárias e Veterinárias, da Universidade Estadual Paulista “Júlio de Mesquita Filho” – Unesp, Câmpus de Jaboticabal pela formação tão ampla na área de melhoramento genético de plantas.

Ao meu orientador Prof. Dr. Antonio Orlando Di Mauro por me aceitar como orientando, por ser um profissional paciente e me permitir fazer um verdadeiro doutorado, aceitando que eu pusesse minhas ideias científicas em prática.

Ao professor Dr. Adhemar Sanches pelos ensinamentos como meu supervisor na disciplina Estatística II (Graduação em administração UNESP/FCAV) e pelas horas dedicadas à minha aprendizagem discutindo sobre métodos Bayesianos.

Aos professores do melhoramento de plantas que atuaram nas disciplinas cursadas por mim: Dr. Alessandro de Mello Varani, Dr. Dilermando Perecin, Dr. Gustavo Vitti Moro, Dr. Rinaldo Cesar de Paula, Dr^a. Janete Aparecida Desidério, Dr^a. Luciana Rossini Pinto e Dr^a. Sandra Helena Uneda-Trevisoli.

Aos professores do melhoramento de animais: Dr. Danisio Prado Munari, Dr. Fernando Sebastian Baldi Rey, Dr. Guilherme Jordão de Magalhães Rosa, Dr. Henrique Nunes de Oliveira e Dr. Roberto Carvalheiro. E ao professor Dr. Carlos Tadeu dos Santos Dias pelos ensinamentos nos tópicos especiais “Métodos Estatísticos Multivariados”.

Aos meus amigos da UNESP/FCAV pela amizade duradoura, Alisson Coutinho, Andrea Andrade, Carlos Eduardo (Carlão), Guilherme Diniz, Gustavo Hugo, Hudson Rabelo, Lucas Santos e Natália Aguilar.

Aos colegas do Laboratório de Biotecnologia pelas trocas de experiências: Amanda Prado, Aretha Corrêa, Éder Groli, Eduardo Bizzari, Elise Pereira, Bruno Henrique, Marcela Marconato, Melina Pirotta, José Arantes, Paulo Selestrino e Wallace Leite.

Aos meus amigos da igreja pelo acolhimento e incentivos constantes: Pr. Alceu, Ângela Santos, Varner, Reginaldo, Max Miller, Helen Roberta, Michéle Santos, Leandro Braz, Mirelle Sara, André Perdiz, Denis Neves, Sirley Santos, Alex, Rodolfo e Isaac.

Aos funcionários do Departamento de produção Vegetal Mônica, Geraldo, Rubens, Sebastião, Mauro e Gabi por desempenharem um papel importante, mantendo o setor funcionando.

Aos funcionários da fazenda de pesquisa da FCAV/UNESP, sobretudo o Marcelo e o tratorista João, que muito contribuíram com o andamento dos experimentos em campo.

A todos que contribuíram de forma direta ou indireta para a realização deste trabalho,

Obrigado!

SUMÁRIO

RESUMO.....	ii
ABSTRACT.....	iii
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA.....	2
2.1 Aspectos gerais e melhoramento da soja	2
2.2 Experimentos multiambientes na cultura da soja.....	5
2.3 Modelos mistos e valores genéticos em soja.....	6
2.4 Abordagem Bayesiana para obtenção de valores genéticos	10
3 MATERIAL E MÉTODOS.....	12
4 RESULTADOS E DISCUSSÃO	15
4.1 Significância das fontes de variação via análise de deviance e seleção das distribuições a priori	15
4.2 Comparação entre os métodos Bayesianos e modelos mistos e desempenho dos genótipos pelos dos ranques das médias	17
4.3 Correlação entre as metodologias Bayesianas, modelos mistos via REML/BLUP, médias ajustadas (lsmeans) e médias aritméticas simples	22
4.5 Parâmetros genéticos para os métodos Bayesiano e modelos mistos	24
5 CONCLUSÕES	26
REFERÊNCIAS.....	27

ABORDAGEM BAYESIANA E MODELOS MISTOS PARA EXPERIMENTOS MULTIAMBIENTES NA CULTURA DA SOJA

RESUMO - A utilização da abordagem Bayesiana pode permitir maior eficiência na aplicação de modelos complexos no melhoramento de plantas, tendo em vista a disponibilidade da computação com alto poder de processamento de dados. Objetivou-se neste estudo a obtenção de valores genéticos pela abordagem Bayesiana, comparando com a inferência frequentista (modelos mistos, lsmeans e médias aritméticas simples), em experimentos multiambientes na cultura da soja. Foram avaliadas 51 linhagens de soja e mais 4 testemunhas no delineamento em blocos casualizados em 6 ambientes com 3 repetições para a característica produtividade de grãos da soja kg.ha^{-1} . Os parâmetros de interesse foram obtidos com o método de Monte Carlo via cadeias de Markov (MCMC) na obtenção de amostras da distribuição a posteriori conjunta. Como informação a priori foram utilizadas as distribuições half-normal a partir dos valores da variância de 18 genótipos de experimentos anteriores e relacionados, bem como a distribuição uniforme. Os valores genotípicos pela abordagem Bayesiana diferiram das médias com a inferência frequentista (modelos mistos, lsmeans e médias aritméticas simples), em experimentos multiambientes na cultura da soja, para os genótipos com alta e baixa produção de grãos. Para os demais casos os quatro métodos testados foram equivalentes. A abordagem Bayesiana com informações a priori derivadas de experimentos anteriores com soja pode ser utilizada nas análises no melhoramento genético dessa cultura. O método de modelos mistos (REML/BLUP) apresentou valores de médias e parâmetros genéticos bem próximos do método Bayesiano para herdabilidade em nível de média (h^2_{mg}), Acurácia na seleção dos genótipos (Acgen), Coeficiente de variação genético (CVgi%) e Coeficiente de variação ambiental (CVe%).

Palavras-chave: *Glycine max*, distribuição a priori no melhoramento genético, modelagem matemática.

BAYESIAN AND MIXED MODELS APPROACH TO MULTIENVIRONMENT TRIALS IN SOYBEAN

ABSTRACT - The Bayesian approach can allow higher efficiency in solving complex models in plant breeding especially with the availability of computing with high-powered data processing. This study aimed to obtain genetic values using Bayesian approach and to compare with frequentist inference (mixed models, lsmeans and simple arithmetic mean) to multienvironment trials in soybean. We evaluated 51 soybean lines and 4 more checks in a randomized complete block design in 6 environments with 3 replications to soybean grain yield kg.ha^{-1} . The parameters of interest were obtained with the method of Monte Carlo Markov chains (MCMC) to obtain samples of the joint posterior distribution. As a priori information the half-normal distribution was used from the values of the variance of 18 genotypes from previous and related experiments as well as the uniform distribution. The genotypic values via Bayesian approach differed from the average with frequentist inference in multienvironment trials in soybean, for genotypes with high and low grain yield. In the other cases the four tested methods were equivalents. The Bayesian approach with priori information derived from previous experiments of soybean can be used in genetic breeding of this crop. The method of mixed models (REML/BLUP) showed slightly different values of means and genetic parameters of the Bayesian method to average heritability (h^2_{mg}), accuracy of genotypes selection (Ac_{gen}), coefficient of genetic variation ($CV_{gi}\%$) and coefficient of environmental variation ($CV_e\%$).

Keywords: *Glycine max*, priori distribution in plant breeding, mathematical modeling.

1 INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) tem grande importância na alimentação humana e animal. O óleo extraído dessa leguminosa é utilizado principalmente com fins industriais. A soja pode ser empregada ainda como fonte de proteína, por seu baixo custo e alto teor de óleo. Na alimentação animal é utilizada na composição de ração e ainda como veículo de medicamentos. Em outras áreas torna-se cada vez maior o uso da soja em produtos industriais como biodiesel, glicerina, sabões, plásticos e tinta para impressão (PAIVA; ALVES; HELENO, 2006). Na safra brasileira de 2016/17 a oleaginosa apresentou uma continuada tendência de crescimento da área plantada, atingindo o percentual de 1,6% em relação à safra passada, totalizando uma produção de 105,56 milhões de toneladas, em uma área de 33,77 milhões hectares, com produtividade média de 3.125 kg.ha⁻¹, 9% a mais que a safra anterior. Portanto o Brasil mantém-se no *ranking* como segundo maior produtor de soja, atrás apenas dos Estados Unidos (COMPANHIA NACIONAL DE ABASTECIMENTO, 2017).

No melhoramento genético da soja uma das maneiras de se avaliar o desempenho de genótipos é por meio de experimentos multiambientes, onde são avaliados genótipos em diferentes locais, safras e anos. Esses experimentos são extensivamente utilizados no melhoramento genético de plantas, em que alguns genótipos são avaliados, sendo feita a mensuração diferencial da variável resposta em cada local, como consequência da interação genótipos com ambientes (G x E).

Em programas de melhoramento de culturas bem estabelecidos, as informações sobre a variância genotípica e ambiental estão disponíveis a partir das análises estatísticas dos testes das cultivares rotineiramente realizados, em que esses parâmetros genéticos podem ser reutilizados para outras análises. A abordagem Bayesiana se concentra em usar tais informações obtidas a partir de ensaios anteriores ou de outros materiais de constituição genotípica similar para utilizar como informação à priori em experimentos futuros (SINGH; AL-YASSIN; OMER, 2015).

Modelagem de dados reais de ensaios envolvendo multiambientes, geralmente, requer computação para modelos complexos como análises conjuntas e multivariadas. Com a disponibilidade de atual nível de poder computacional, a

abordagem Bayesiana torna-se exequível para lidar com complexos modelos estatísticos (OMER et al., 2015).

No melhoramento genético da soja as análises estatísticas multiambientais têm sido feitas do ponto de vista frequentista e, portanto, não utilizam dados de experimentos anteriores como informação a priori. A abordagem frequentista é restritiva e utiliza modelos em que os parâmetros lineares e parâmetros de variância são tratados como constantes e ignoram qualquer informação a priori sobre esses parâmetros. Em contrapartida, a abordagem Bayesiana integra a informação prévia disponível a partir de estudos anteriores com o conjunto de dados atual e fornece uma cobertura mais realista e mais ampla para inferência estatística (SINGH; AL-YASSIN; OMER, 2015).

Na cultura da soja existem relatos de trabalhos em que a metodologia Bayesiana foi utilizada para definir a formação de cluster (PRIOLLI et al., 2013), cálculos de mudança fenológica (SHEN; LIU, 2015) e redes Bayesianas para estudar relação entre caracteres (VALENTIM; SILVA; ALVES, 2007). Contudo, até o presente momento no melhoramento genético de soja com dados fenotípicos não existe utilização de métodos Bayesianos para obtenção de valores genotípicos.

A predição dos valores genotípicos dos materiais superiores é o problema central do melhoramento de plantas, pois é necessário conhecer os valores verdadeiros dos componentes de variância, tornando-se imprescindível o uso de métodos adequados (BORGES et al., 2010). Nesse contexto, os métodos Bayesianos apresentam-se como uma análise alternativa na estimação de valores genotípicos podendo ser utilizado na área de melhoramento de plantas.

Objetivou-se neste estudo a obtenção das estimativas de parâmetros genéticos pela metodologia Bayesiana, comparando com os resultados obtidos pela inferência frequentista (modelos mistos, *lsmeans* e médias aritméticas simples), em experimentos multiambientais na cultura da soja.

2 REVISÃO DE LITERATURA

2.1 Aspectos gerais e melhoramento da soja

A soja é uma leguminosa anual, pertencente à classe *Dicotyledoneae*, família *Fabaceae*, gênero *Glycine* e espécie *Glycine max* (L.) Merrill. É uma planta de

origem oriental, que é consumida em larga escala nos países asiáticos, sob as mais diversas formas. Esta cultura é considerada milenar e constitui-se em uma das principais oleaginosas cultivadas mundialmente (PAIVA; ALVES; HELENO, 2006).

O melhoramento genético da soja é um processo contínuo de desenvolvimento de novas cultivares. Os programas de melhoramento possuem objetivos gerais e específicos e buscam a solução das limitações reais ou potenciais das cultivares frente aos fatores bióticos e abióticos que interferem na produção da cultura. Grande parte da variabilidade genética desta cultura é mantida e conservada em Bancos de Germoplasma existentes em vários países orientais e ocidentais (ALMEIDA et al., 1999). A soja é a espécie em que se conseguiu aproximar mais da maximização da produtividade, por meio da associação do uso racional e eficiente do ambiente com o melhoramento genético. O ponto relevante do melhoramento é o fato de o uso de novas cultivares não representar custos adicionais aos agricultores (MAIA et al., 2013).

Com relação ao cultivo comercial da soja, atualmente os campos de produção dividem-se em campos de cultivares denominado convencional e campos de cultivares denominada transgênica. Por definição, os transgênicos são organismos ou culturas geneticamente modificados (OGMs) que contêm um gene que foi artificialmente inserido. Nas culturas convencionais a transferência genética é naturalmente feita por polinização (ALMEIDA; LAMOUNIER, 2011).

As características de primeira geração introduzidas em cultivares de soja, por meio da biotecnologia, foram resistência a herbicidas, como a resistência ao glifosato sendo o mais comum. No futuro, novas sojas resistentes a herbicidas são prováveis. Contudo, é importante ressaltar que a biotecnologia moderna, em si, nunca irá substituir a pesquisa melhoramento de plantas, mas sim melhorar e aperfeiçoar sua eficiência (SLEPER; SHANNON, 2003).

No melhoramento genético desta cultura estão envolvidos cruzamentos ou híbridos de polinização manual seguido pela seleção, testes, e, finalmente, lançamento de uma cultivar superior (SLEPER; SHANNON, 2003). Híbridações são realizadas para desenvolver germoplasma com variabilidade genética e as populações segregantes são conduzidas, geralmente por métodos tradicionais de melhoramento de plantas autógamas, para permitir a seleção e a avaliação de genótipos com as características agrônômicas desejadas nas novas cultivares (ALMEIDA et al., 1999).

Contudo, como um membro da família Fabaceae, a soja apresenta pseudocleistogamia rigorosa, onde a polinização ocorre com o botão floral fechado. Esta característica pode ter um forte impacto na manutenção da homogeneidade e redução na variação genômica, para a espécie, a qual pode ter sido ainda mais acentuada pelo processo de domesticação (LAM et al., 2010). Isto reflete, atualmente, no desenvolvimento de novas cultivares.

Diversos programas de melhoramento genético contribuíram para o desenvolvimento de cultivares de soja de alto rendimento e adaptadas às diferentes condições agroclimáticas do Brasil (PRIOLLI et al., 2004) e do mundo.

As pesquisas em melhoramento genético têm contribuído significativamente para o desenvolvimento e garantia do sucesso da cultura da soja no Brasil.

Os ganhos de produção por área, produtividade, associada à ampliação das fronteiras agrícolas, especialmente para as áreas de cerrados, são contribuições do melhoramento, pela obtenção de cultivares mais produtivas e adaptadas em diferentes ambientes (ROSSMANN, 2001).

Os crescimentos da produção e produtividade da soja brasileira estão também associados aos avanços tecnológicos, científicos como à disponibilidade de tecnologias do setor produtivo. Estima-se que a metade do incremento da produtividade das principais cultivares agrônômicas nos últimos cinquenta anos seja mérito do melhoramento genético de plantas (BORÉM; CONDORI; MIRANDA, 2008).

Em se tratando de soja, um aspecto de grande importância para o melhorista é conhecer como um genótipo se desenvolve em cada região e, para tal, são realizados experimentos, os quais fornecerão diretrizes para o estudo e/ou desenvolvimento de novas cultivares de soja (SILVEIRA; DI MAURO; CENTURION, 2008).

O melhorista busca ganhos genéticos significativos, sendo este um ponto fundamental para o desenvolvimento de novas cultivares. Dessa forma, as características desejáveis em cultivares de soja são: maior resistência às doenças limitantes, insetos pragas e aos nematoides associados à cultura; boa resistência ao acamamento e à deiscência precoce; boa qualidade fisiológica da semente; adaptação às condições locais de ambiente e ideótipo de planta adequada ao sistema agrícola utilizado na região produtora, além de precocidade (BACAXIXI et al., 2011).

A cultura da soja vem crescendo no país, onde está relacionada aos avanços

científicos e a disponibilização de tecnologias ao setor produtivo. A mecanização, o melhoramento genético na criação de cultivares altamente produtivas e adaptadas às diversas regiões, além da identificação e solução para os principais fatores responsáveis por perdas no processo de colheita, são fatores desse avanço (FREITAS, 2011).

2.2 Experimentos multiambientes na cultura da soja

O cultivo extensivo da soja, contemplando ambientes diversificados, provoca uma resposta diferencial dos genótipos. Neste sentido, a GE (interação de genótipo com ambiente) representa aspecto relevante no contexto do melhoramento (MAIA et al., 2013). A presença da interação complexa também reduz a correlação entre os valores genotípicos e fenotípicos, diminuindo os ganhos genéticos com a seleção e complicando o trabalho dos melhoristas (NUNES, 2000).

Para a maioria das características de importância econômica, em especial a produtividade de grãos, o fenótipo de soja a ser selecionado depende do genótipo, do ambiente e da interação do genótipo x ambiente. Essa interação ocorre devido à falta de estabilidade dos genótipos nos vários ambientes, refletindo nas diferentes respostas dos genótipos às mudanças ambientais. Considerando as inúmeras variações ambientais em que a soja é comumente submetida no Brasil, é esperado que a interação genótipo x ambiente assuma papel fundamental na manifestação fenotípica, devendo, portanto, ser estimada e considerada no programa de melhoramento genético e na indicação de cultivares (PRADO et al., 2001).

O comportamento de genótipos pode ser elucidado pelo estudo GE com sua partição em parâmetros de adaptabilidade (responsividade ao estímulo ambiental) e estabilidade produtiva (previsibilidade de comportamento). Em particular, a estabilidade específica de genótipos a multiambientes possibilita tirar proveito desse efeito estando, via de regra, associado a elevadas produtividades. Pela mesma razão, sob o ponto de vista de recursos genéticos, a exploração dessa interação feita por meio do zoneamento ecológico é interessante para manter a variabilidade genética da espécie, principalmente, porque a soja cultivada possui base genética estreita. Contudo, para que seja possível tirar proveito desses efeitos positivos, é preciso dispor de métodos estatísticos adequados para se estimar e explorar a

interação, permitindo assim recomendações regionalizadas (DUARTE; ZIMMERMANN, 1995).

Em relação à adaptação produtiva de soja, em um estudo procurando encontrar locais chave para a condução de programa de melhoramento, surgiu o termo “zonas agronômicas” para recomendação de cultivares de soja. Além disso, o fato de que a otimização da produção onde nenhum genótipo é o melhor em todas as localidades está ligada à possibilidade de que a rede experimental possa ser subdividida em extratos relativamente homogêneos do ponto de vista da interação genótipos x ambientes. Portanto, tais extratos ou zonas agronômicas, são áreas não necessariamente contínuas de uma região de cultivo, mas que apresentam um conjunto de condições ambientais que possibilitam aos mesmos genótipos serem sempre os melhores nesse conjunto de locais (PACHECO et al., 2007).

Em virtude da existência da interação cultivares x ambientes, torna-se pouco provável a superioridade de uma cultivar de soja em todos os ambientes. Neste sentido, em programas de melhoramento genético de soja é necessário avaliar a similaridade de ambientes e determinar o número de ambientes necessários para avaliação de cultivares. Uma vez que o principal objetivo de um programa de melhoramento é selecionar genótipos com produtividade elevada e consistente, nos mais diversos ambientes (PELUZIO et al., 2012).

Quando o comportamento de duas cultivares são concordantes em dois ambientes distintos, a interação é chamada de interação simples, não acarretando maiores problemas. Entretanto, quando as cultivares possuem comportamento diverso, a interação é denominada complexa. Considerando um número maior de ambientes e de cultivares, a presença de interação complexa quase sempre indica a existência de cultivares especificamente adaptadas a ambientes particulares, bem como de outras com adaptação (RAMALHO; SANTOS; ZIMMERMANN, 1993).

2.3 Modelos mistos e valores genéticos em soja

O problema central do melhoramento genético é a predição dos valores genotípicos dos vários candidatos à seleção. Esta predição necessita de componentes de variância conhecidos ou estimados com precisão (RESENDE, 2004). Anteriormente, pelo método dos quadrados mínimos, as covariâncias foram estimadas e interpretadas em termos da sua esperança matemática, resultando em

componentes de variância. Para isso, algumas suposições tiveram que ser feitas: 1) modelo de aditividade 2) distribuição normal dos dados, independência e homogeneidade dos erros. Atualmente, os componentes de variância e as variâncias dos efeitos aleatórios em modelos lineares mistos podem ser estimados diretamente no conjunto de dados, em que não é necessário para atender às premissas acima (CARVALHO; FRITSCH NETO; GERALDI, 2008).

Quando o modelo apresenta todos os demais efeitos fixos ele é denominado de modelo linear fixo; quando o modelo linear apresenta, além do erro experimental, outros efeitos aleatórios em comum com outros efeitos fixos, além da média, são denominados de modelo misto. Nos modelos lineares mistos, o êxito da modelagem está diretamente ligado às estruturas de variâncias e covariâncias das variáveis aleatórias que são explicadas por fatores aleatórios, após a eliminação dos efeitos fixos, e da caracterização dos efeitos residuais. Os modelos lineares, nos parâmetros, possuem pelo menos um efeito aleatório designado por erro experimental (PEREIRA; GRAVINA; THIÉBAUT, 2012).

O modelo misto de seleção foi apresentado por Henderson, em 1973, porém, muito antes havia sido concebido por ele mesmo por volta de 1949, quando era aluno da disciplina de Matemática e estatística, ministrada por Mood, ao resolver um problema estatístico simples, utilizando a melhor predição linear não viesada (BLUP). No BLUP, a média é a mesma para todos os tratamentos genéticos, uma vez que, em termos de processo de estimação, se trata de uma mesma população (GONÇALVES; FRITSCH NETO, 2012).

Por sua vez, o procedimento da Máxima Verossimilhança Restrita (REML) foi criado pelos pesquisadores ingleses Desmond Patterson e Robin Thompson em 1971, e hoje se constitui na metodologia padrão para a análise estatística em uma grande gama de aplicações. O REML é uma generalização da ANOVA para situações mais complexas. Para situações simples, os dois procedimentos são equivalentes, mas para as situações mais complexas encontradas na prática, a ANOVA é um procedimento apenas aproximado (RESENDE, 2007b).

Tem-se, assim, o procedimento ótimo de predição de valores genéticos, como sendo o BLUP, e o procedimento ótimo de estimação de componentes de variância, como sendo o REML. As principais vantagens práticas do REML/BLUP (máxima verossimilhança restrita / melhor predição linear não viesada) são: permite comparar indivíduos ou cultivares ao longo do tempo (gerações e anos) e espaço (locais e

blocos); permite a simultânea correção para os efeitos ambientais, estimação de componentes de variância e predição de valores genéticos; permite lidar com estruturas complexas de dados (medidas repetidas, diferentes anos, locais e delineamentos); pode ser aplicado a dados desbalanceados e a delineamentos não ortogonais. No caso de dados desbalanceados, a ANOVA conduz a imprecisas estimativas de componentes de variância e conseqüentemente a predições não acuradas de valores genéticos (RESENDE, 2007a).

Além disso, a utilização de modelos mistos para estimação de parâmetros melhora o entendimento de mecanismos genéticos, é necessária para predição de valores genotípicos (índice de seleção/ BLUP) e para aperfeiçoar programas de melhoramento e a predição de respostas (GONÇALVES; FRITSCH NETO, 2012).

Os preditores BLUP e estimadores REML são usados nas várias situações associadas aos diferentes delineamentos experimentais tais como blocos ao acaso, látice, linha e coluna e blocos aumentados. São usados também associados aos vários delineamentos de cruzamentos, tais como progênies de polinização aberta, cruzamentos dialélicos, fatoriais, hierárquicos e testes clonais. O BLUP também tem sido importante na predição de híbridos simples de milho e de outras espécies. Os delineamentos mais complexos como os dialelos e fatoriais permitem uma grande gama de inferências, tais quais aquelas sobre os efeitos aditivos (capacidade geral de combinação), efeitos de dominância e da capacidade específica de combinação bem como suas variâncias, tanto em nível individual quanto de famílias (RESENDE, 2004).

Também, os delineamentos de cruzamentos associados à autofecundações podem ser analisados via REML/BLUP. Em espécies autógamas é comum a avaliação, em várias gerações de autofecundação, de linhagens obtidas a partir de cruzamentos entre dois genitores divergentes. Em espécies de reprodução vegetativa (como a cana-de-açúcar e espécies forrageiras) o procedimento BLUP é o ideal na seleção de indivíduos para clonagem na fase inicial do melhoramento, considerando simultaneamente as informações do indivíduo, da família, do delineamento experimental e do parentesco entre famílias e genitores (RESENDE, 2007a).

No caso da análise de modelos mistos com dados desbalanceados, os efeitos do modelo não são testados via testes F, tal como se faz no método de análise de variância. Nessa situação, para os efeitos aleatórios, o teste cientificamente

recomendado é o teste da razão de verossimilhança (LTR). Para os efeitos fixos, um teste F aproximado pode ser usado. Um quadro similar ao quadro de análise de variância pode ser elaborado, sendo denominada análise de deviance (ANADEV), e é estabelecido segundo os seguintes passos (RESENDE, 2007b):

- a) Obtenção do logaritmo do ponto de máximo da função de verossimilhança residual (L) para modelos com e sem o efeito a ser testado;
- b) Obtenção da deviance $D = -2 \log L$ para modelos com e sem o efeito a ser testado;
- c) Obtenção da diferença entre as deviances para os modelos sem e com o efeito a ser testado, alcançando a razão de verossimilhança (LR);
- d) Teste via LTR da significância dessa diferença usando o teste qui-quadrado com 1 grau de liberdade.

Em suma, no melhoramento genético de plantas, os modelos mistos podem ser aplicados na análise de experimentos, sendo obtidas as seguintes propriedades: maximização da acurácia seletiva, minimização do erro de predição, predição não viciada de valores genéticos, maximização do ganho genético por ciclo de seleção, maximização da probabilidade de selecionar o melhor entre dois genótipos; maximização da probabilidade de selecionar o melhor entre vários genótipos. É importante relatar, ainda, que o BLUP é o mais eficiente índice de seleção em termos de uso das informações de parentes, pois usa todas as informações disponíveis, que são os efeitos aleatórios do modelo estatístico (RESENDE, 2004).

Em progênies de soja a metodologia modelos mistos faz a predição de valores genéticos dos cruzamentos entre genótipos visando selecionar progênies de soja que contem alto teor de óleo e/ou proteína e alta produtividade, sendo a metodologia REML/BLUP eficaz na determinação de melhores cruzamentos (PINHEIRO et al., 2013).

Andrade et al. (2016) selecionaram genótipos de soja oriundos de cruzamentos entre linhagens convencionais e transgênicas (*Roundup Ready*), utilizando, conjuntamente, as abordagens REML/BLUP, análise de fatores e componentes principais, para as variáveis número de dias para florescimento (NDF), altura da planta no florescimento (APF), número de dias para a maturidade (NDM), altura da planta na maturidade (APM), Altura de inserção da primeira vagem (AIV), acamamento (AC), valor agrônômico (VA), número de ramos por planta (NR), número de nós por planta (NN), número de vagens por planta (NV), produção de

grãos (PG) e peso de cem sementes (PCS), tal combinação é eficiente para a seleção de genótipos de soja com bons componentes de produção, mais precoces, de menores portes e com resistência ao acamamento.

A metodologia REML/BLUP pode ser utilizada em ensaios de competição final de genótipos de soja para predição de valores genotípicos de genótipos pertencentes aos três grupos de maturação (superprecoce, precoce e semiprecoce), podendo ser feito o uso da informação de parentesco obtida por genealogia ou por uso da informação de parentesco obtida por marcadores moleculares microssatélites aumentando a precisão das predições de valores genotípicos (SSR) (GODOI, 2014).

Ainda em ensaios finais dessa cultura, o método MHPRVG (média harmônica da performance relativa dos valores genotípicos), calculado via modelos mistos pode ser utilizado para determinar valores genéticos de linhagens de soja mais produtivas com maior estabilidade e adaptabilidade (GOMEZ, 2014; RESENDE, 2007b).

2.4 Abordagem Bayesiana para obtenção de valores genéticos

Na abordagem Bayesiana todos os efeitos do modelo, sejam fixos ou aleatórios conforme definidos sob o ponto de vista frequentista, são tratados como aleatórios. Em outras palavras, as definições de efeitos fixos e aleatórios na abordagem Bayesiana não surgem naturalmente como na abordagem frequentista (SORENSEN; GIANOLA, 2002).

A abordagem clássica ou frequentista diferencia-se da abordagem Bayesiana, pois os parâmetros são vistos como variáveis aleatórias cujo comportamento é regulado por uma distribuição de probabilidade que se assume sobre seus possíveis valores, traduzindo uma informação inicial (ou a priori) que se tenha sobre os parâmetros, antes mesmo de se obter os dados. A seguir, essa informação a priori é combinada com aquela proveniente dos dados, produzindo uma informação a posteriori sobre os parâmetros. Essa informação a posteriori se expressa na chamada distribuição a posteriori dos parâmetros, que é utilizada para se fazer inferência sobre os mesmos (COSTA; SANCHES; MUNARI, 2009).

Embora tendo utilidade notória no melhoramento genético, a abordagem clássica de modelos mistos apresenta algumas limitações que podem ser contornados pela inferência Bayesiana, que consiste em uma abordagem estatística que trata os parâmetros populacionais como variáveis aleatórias e que permite obter

intervalos de confiança que são denominados de intervalos de credibilidade, fazendo uso de métodos de simulação de Monte Carlo. Ou seja, trata-se de uma excelente alternativa quando, na inferência clássica, a distribuição de amostragem de um estimador for complexa, inviabilizando a construção de intervalos de confiança (NOGUEIRA et al., 2003).

Recentemente, métodos Bayesianos vêm sendo utilizados como uma opção para solucionar problemas relacionados à avaliação de mérito genético em populações animais. A aplicação de métodos Markov Chain Monte Carlo (MCMC), dentre os quais se destaca a amostragem de Gibbs, pode ser utilizada de forma a propiciar uma inferência Bayesiana. O método Gibbs Sampling (GS) é uma técnica de integração numérica por simulação, muito usual em situações nas quais a integração analítica completa é impossível. O GS é aplicável à estimação de componentes de variância e predição dos valores genéticos, permitindo, por suas propriedades, a inferência Bayesiana (CARNEIRO JÚNIOR et al., 2010).

Em um experimento com dados reais de animais da raça Guzerá, compararam a aplicação da análise Bayesiana na estimação de componentes de variância e covariância e na predição dos valores genéticos, com os resultados obtidos pela metodologia REML. A análise Bayesiana apresentou grande flexibilidade tanto em relação aos modelos que podem ser utilizados nas análises quanto às inferências que podem ser realizadas a partir dos resultados (OLIVEIRA; LÔBO; PEREIRA, 2000).

Comparando o método E-BLUP (Empirical BLUP) utilizando modelos mistos com o método Bayesiano, os dois métodos apresentam resultados semelhantes, quando distribuições a priori não informativas são utilizadas, e as populações de tamanho grande, de modo geral, apresentaram melhores predições de valores genéticos. Na metodologia Bayesiana, o aumento no nível de informação a priori influencia positivamente as predições dos valores genéticos, principalmente para as populações pequenas. O método Bayesiano é indicado para populações de tamanho pequeno quando há disponibilidade de distribuições a priori informativas (CARNEIRO JÚNIOR et al., 2010).

Avaliando modelos mistos na seleção de famílias de cana de açúcar aparentadas sob o enfoque clássico e Bayesiano, foram verificadas as variâncias

genética e residual, a herdabilidade e os valores genotípicos das famílias, em função de seus intervalos de confiança ou de credibilidade. Foi verificado que o procedimento Bayesiano foi mais eficiente quando se consideraram, no modelo, a informação de parentesco e a distribuição a priori como informativa. Nos demais casos os dois procedimentos apresentaram resultados equivalentes. Além disso, foi verificado que na análise Bayesiana, a inclusão da distribuição a priori informativa leva a intervalos de credibilidade mais estreitos. De acordo com os intervalos de credibilidade, na maioria dos casos avaliados (80%), a informação do parentesco utilizada aumenta os valores genotípicos preditos e reduz as estimativas dos componentes de variâncias aumentando a herdabilidade (SILVA et al., 2013b).

A predição de valores genéticos e componentes de variância via Inferência Bayesiana ainda são raras no melhoramento de plantas. Desse modo, estudos futuros devem trabalhar para implementar uma estrutura Bayesiana para análise de modelos mistos aplicada ao melhoramento e explorar diferentes possibilidades no uso de informações a priori (SILVA et al., 2013a). Além disso, o mesmo autor relata que em milho pipoca no primeiro ciclo de seleção, a utilização de priores informativas por meio de meta-análise fornece resultados claramente distintos em relação ao uso de priores não informativas apenas para a produção de grãos. As diferenças entre as análises são restritas aos componentes de variância e herdabilidade.

Para a cultura da soja não foram encontradas publicações com a utilização da metodologia Bayesiana utilizando informações a priori com o intuito de estimar os componentes de variância e valores genéticos.

3 MATERIAL E MÉTODOS

Foi avaliada a variável produtividade de grãos de soja em kg.ha^{-1} para linhagens avançadas do Programa de Melhoramento de Soja da Universidade Estadual Paulista. O delineamento utilizado foi em blocos completos casualizados (DBC) com três repetições. Foram avaliadas 51 linhagens de soja originadas de cruzamentos simples, duplos, quádruplos e óctuplos a partir de diferentes genitores. Além disso, nos experimentos foram utilizadas as cultivares testemunhas CD-216, Conquista (MG-BR46), Potência e V-Max.

Os ensaios foram realizados nas seguintes safras e localidades no estado de São Paulo: Pindorama (Safr 2013/2014), Jaboticabal (2013/2014), Jaboticabal (Safr 2014/2015), Jaboticabal (Safr 2015/2016), Piracicaba Safr (2013/2014) e Piracicaba (Safr 2014/2015), totalizando 6 ambientes. Os fatores safras e locais foram combinados e denominados de ambientes. Portanto, todas as inferências foram feitas sem particionar safras e locais conforme recomendação da literatura (OMER et al., 2015).

Pela abordagem frequentista, o modelo para produtividade nos múltiplos experimentos conduzidos, em termos de efeitos do ambiente, blocos dentro de ambientes, genótipos e genótipos x ambientes pode ser descrito como:

$$Y_{ijk} = \mu + E_j + R_{k(j)} + G_i + (GE_{ij}) + e_{ijk} \quad (1)$$

Onde:

Y_{ijk} = produtividade para o i -ésimo genótipo no k -ésimo bloco do j -ésimo ambiente, μ é a média geral, E_j é o efeito do j -ésimo ambiente, $R_{k(j)}$ é o efeito do k -ésimo bloco no ambiente j , G_i é o efeito do i -ésimo genótipo, (GE_{ij}) é a interação entre o i -ésimo genótipo e o j -ésimo ambiente, e e_{ijk} é o erro aleatório. Os ambientes foram considerados fixos e repetição dentro de ambientes, genótipos, interação genótipos por ambientes e erro considerados como aleatórios.

Em geral, além da μ , os seguintes pressupostos de efeitos aleatórios foram feitos:

$$R_{k(j)} \sim N(0, \sigma_R^2), G_i \sim N(0, \sigma_G^2), GE_{ij} \sim N(0, \sigma_{GE}^2) \text{ e } e_{ijk} \sim N(0, \sigma_e^2)$$

Onde $\sim$ significa “é distribuída como” e $N(0, \sigma^2)$ é a distribuição normal com média zero e variância σ^2 .

Na abordagem Bayesiana, os componentes de variância dos efeitos e interações na equação foram assumidos como sendo variáveis aleatórias com distribuições a priori com parâmetros conhecidos. Tais distribuições a priori, como half-normal e uniforme, têm sido estudadas e recomendadas para os componentes de variância ou para os componentes correspondentes ao desvio padrão ou parâmetro de escala (GELMAN, 2006).

Os parâmetros estudados e distribuição a posteriori foram obtidos com o método de Monte Carlo via cadeias de Markov (MCMC), com o amostrador de Gibbs. A análise computacional foi composta de 3 cadeias, cada uma com 10.000 iterações (as primeiras 5.000 descartadas), com espaçamento dos pontos amostrados de 15 (*thinning*), e número de simulações igual a 1002 iterações salvas. Como o Gibbs sampler é um algoritmo iterativo, faz-se necessário a verificação de sua convergência pela estatística de Gelman-Rubin ($\hat{R}$) que mensura a razão da variância entre cadeias sobre a variância dentro de cadeias. Verifica-se que todos os valores do $\hat{R}$ para genótipos e ambientes ficaram abaixo de 1,1 confirmando a convergência das iterações (GELMAN; RUBIN, 1992).

Foram testados dois tipos de distribuição a priori no presente trabalho. Como tipo 1 foi utilizada a distribuição “half – normal” positiva para as variâncias σ_E^2 , σ_G^2 , σ_R^2 e σ_{GE}^2 , em que os valores para esta distribuição a priori foram calculados com base no inverso das variâncias do ensaio anterior com 18 genótipos de soja que fazem parte dos 51 genótipos avaliados, que são os genótipos: Jab 4, Jab 5, Jab 6, Jab 7, Jab 8, Jab 10, Jab 12, Jab 15, Jab 16, Jab 17, Jab 19, Jab 20, Jab 22, Jab 23, Jab 24, Jab 39, Jab 40 e Jab 41. Os dados dos 18 genótipos que estão no presente trabalho foram avaliados anteriormente em três experimentos conduzidos no estado de São Paulo, nos municípios de Jaboticabal (nos anos agrícolas 2011/2012 e 2012/2013) e Piracicaba (ano agrícola 2012/2013) (GOMEZ, 2014).

Como priori tipo 2 foi testada a distribuição uniforme no intervalo [0, 5.000.000], esses valores foi escolhido com base no trabalho relacionado ao presente estudo (GOMEZ, 2014). A Seleção das priores consideradas mais adequadas para análise Bayesiana foi realizada utilizando as estatísticas: -2 x log-likelihood para as médias a posteriori dos parâmetros, média posteriori de -2 x log-likelihood, número efetivo de parâmetros e critério de informação da deviance (DIC) (OMER et al., 2014a).

A distribuição half-normal vem se apresentando como uma das mais utilizadas para estudos com experimentação e melhoramento genético de plantas por apresentar o menor critério de informação da deviance (DIC) para determinação do número de repetições em ensaios de culturas (OMER et al., 2014b) e estimação

de herdabilidade e ganho genético por meio da abordagem Bayesiana (OMER et al., 2014a).

Foram calculadas as médias para os genótipos considerando os métodos Bayesianos, modelos mistos, médias ajustadas (*lsmeans*) e médias aritméticas simples. Foram realizadas as correlações entre as médias e os ranques das médias dos métodos avaliados. A correlação de Pearson foi feita entre médias dos genótipos e a correlação de Spearman calculada entre os ranques.

Todas as análises Bayesianas foram realizadas no software R (R CORE TEAM, 2016) no pacote R2WinBUGS (STURTZ; LIGGES; GELMAN, 2005). O pacote lme4 foi utilizado para cálculos dos BLUPs (valores genotípicos com o método REML usando como matriz de covariância a simetria composta) e variâncias (BATES et al., 2014), e os cálculos da análise de deviance foram realizadas no pacote lme4 (ZEILEIS; HOTHORN, 2002).

4 RESULTADOS E DISCUSSÃO

4.1 Significância das fontes de variação via análise de deviance e seleção das distribuições a priori

Considerando ambiente como único efeito fixo no modelo, uma vez que foram testados três ambientes, todas as fontes de variação foram estatisticamente diferentes, com significância menor que 1% pelos testes F e LRT (Tabela 1).

Tabela 1. Análise de deviance para as fontes de variação aleatórias e teste F para ambientes, considerando produtividade de grãos kg.ha⁻¹, Jaboticabal, SP, 2016

Fontes de variação	Deviance	Teste estatístico	P-valor
Ambientes	(†)	F=29,38	<0,00001
Rep/Ambientes	15935,10 ⁺	LRT=10,22	0,001392
Genótipos	15937,48 ⁺	LRT=12,59	0,000388
Genótipos*Ambiente	15932,09 ⁺	LRT=7,20	0,007292
Modelo completo	15924,89		

*†. Valor não calculado para efeitos fixos, F. Valor F calculado, LRT. Teste da razão da verossimilhança pelo qui-quadrado com um grau de liberdade. Qui-quadrado tabelado a 1%: 3,83 e a 5%: 6,63. +. Deviance do modelo ajustado sem o referido efeito.

Como houve parcelas perdidas em alguns dos experimentos, o mesmo é considerado desbalanceado e, portanto, os fatores aleatórios devem ser testados com a análise da deviance. Na análise de experimentos com dados desbalanceados os efeitos aleatórios em modelos mistos não são testados via testes F tal como se faz no método da análise de variância. Nesse caso, para os efeitos aleatórios, o teste cientificamente recomendado é o teste da razão de verossimilhança (LRT) que utiliza o teste do qui-quadrado com um grau de liberdade (RESENDE, 2007b).

A distribuição a priori mais adequada para os cálculos dos valores genéticos, utilizando o método MCMC, foi a half-normal positiva que apresentou menor critério de informação da deviance (DIC), número efetivo de parâmetros (P_D), deviance das médias a posteriori dos parâmetros ($\hat{D}$) e média posteriori da deviance ($\bar{D}$) (Tabela 2). O valor do DIC da distribuição half-normal foi menor que o valor da distribuição uniforme. Uma vez que o DIC é uma estimativa do erro predito esperado, quanto menor o DIC mais o modelo estão bem ajustado. Isso indica, portanto que distribuição a priori half-normal deve ser utilizada ao invés da distribuição uniforme. Portanto, todos os componentes de variância foram calculados utilizando a half-normal como distribuição a priori. Ressalta-se ainda que a priori escolhida para cálculos posteriores utilizando o método Bayesiano foi construída com base nos valores do inverso das variâncias de 18 genótipos dos 55 do presente trabalho. No que se refere à avaliação em experimentos genéticos, o uso de informações a priori de experimentos anteriores, conduzem a melhores estimativas das variâncias, diminuindo a variância residual do experimento (CARNEIRO JÚNIOR et al., 2005).

Tabela 2. Estatísticas de discrepância para seleção das priores dos ensaios de soja em seis ambientes, considerando produtividade de grãos kg.ha⁻¹. Jaboticabal, SP, 2016

Modelo de distribuição a priori	$\bar{D}$	$\hat{D}$	P_D	DIC
P1	14981,80	14791,60	190,24	15172,00
P2	15951,60	15901,30	50,34	16001,90

Onde: $\bar{D}$. Média posteriori de $-2 \times \log\text{-likelihood}$. $\hat{D}$. $-2 \times \log\text{-likelihood}$ para as médias a posteriori dos parâmetros. P_D . número efetivo de parâmetros. DIC. Critério de informação da deviance.

P1: $\sigma_E^2, \sigma_G^2, \sigma_R^2$ e σ_{GE}^2 com distribuição independentemente $\sim$ half-normal positiva com $\sigma_E^2 \sim (0, 0,00028)$, $\sigma_G^2 \sim (0, 0,00304)$, $\sigma_R^2 \sim (0, 0,00001)$ e $\sigma_{GE}^2 \sim (0, 0,00170)$

P2: $\sigma_E^2, \sigma_G^2, \sigma_R^2$ e σ_{GE}^2 Independentemente $\sim$ uniforme $(0, 5.000.000)$.

4.2 Comparação entre os métodos Bayesianos e modelos mistos e desempenho dos genótipos pelos dos ranques das médias

O desempenho dos genótipos por meio da abordagem Bayesiana também foi dado com base na média a posteriori e em termos de percentis dos valores previstos e de seus ranques. A média brasileira da produtividade de soja é de 2.870 kg.ha⁻¹ (COMPANHIA NACIONAL DE ABASTECIMENTO, 2017). Considerando os intervalos de credibilidade e o quantil 0,05, oito genótipos: Jab 42, Jab 44, Jab 32, Jab 31, Conquista, Jab 27, Jab 41 e Jab 5 são considerado produtivos, pois apresentam valores do intervalo de credibilidade dentro da produtividade média brasileira de soja. Além disso, a linhagem Jab 42 que está em fase de melhoramento apresentou produtividade maior que a média brasileira (Tabela 3).

Tabela 3. Predição dos valores das médias dos genótipos ordenadas pela média a posteriori e seus ranques pela metodologia Bayesiana e análises frequentistas com predição dos valores genotípicos (utilizando BLUPS), média ajustada e média aritmética para 55 linhagens de soja, considerando produtividade de grãos kg.ha⁻¹. Jaboticabal, SP, 2016.

Ranque da média a posteriori	Genótipos	Média a posteriori	⁽¹⁾ IC 95%		Ranque Médio da posteriori	Ranque a posteriori (Quantis)					⁽²⁾ Valor genotípico (u+g+gem)	Média fenotípica (<i>lsmeans</i>)	Média aritmética	⁽³⁾ N
			Inferi or	Super ior		0,025	0,05	0,5	0,95	0,975				
1	Jab 42	3331,33	2914	3843	1,07	1	1	1	1	2	3514,27	4179,80	4179,80	18
2	Jab 44	2743,28	2390	3168	7,78	2	2	5	22,48	27	2802,50	3057,61	3057,61	18
3	Jab 32	2655,86	2262	3052	11,06	2	2	9	29	34	2707,58	2907,95	2907,95	18
4	Jab 31	2622,42	2260	2994	12,39	2	2	10	32	36	2670,94	2850,19	2850,19	18
5	Conquista	2620,53	2255	3004	12,56	2	2	10	32	36	2666,00	2842,40	2842,40	18
6	Jab 27	2619,51	2225	2991	12,72	2	2	10	33	36,98	2670,05	2848,78	2848,78	18
7	Jab 41	2594,81	2180	2970	13,81	2	2	11	34	38,49	2652,07	2907,69	2827,45	17
8	Jab 5	2590,56	2205	2986	14,03	2	2	11	35	39,98	2636,43	2795,77	2795,77	18
9	Jab 30	2567,34	2159	2934	15,25	2	3	13	37	40	2594,21	2723,61	2820,76	17
10	Jab 45	2563,80	2213	2976	15,59	2	3	14	36	39	2606,41	2737,20	2835,76	17
11	Jab 39	2527,19	2189	2955	17,39	2,03	3	15	39	42	2565,91	2684,60	2684,60	18
12	Jab 1	2506,70	2099	2856	18,49	3	3,5	16	40	43	2557,47	2671,29	2671,29	18
13	Jab 7	2506,69	2085	2857	18,53	3	4	17	41	44	2539,33	2642,69	2642,69	18
14	Jab 43	2498,92	2090	2913	19,14	2	3	17	41	45	2530,52	2597,47	2651,90	17
15	Jab 35	2489,26	2088	2862	19,47	3	4	18	41	44	2503,67	2540,52	2609,15	17
16	Potência	2486,78	2126	2876	19,60	3	4	18	42	45	2513,25	2601,56	2601,56	18
17	Jab 2	2439,57	2022	2813	22,49	4	5	21	44	47	2452,12	2488,70	2446,42	17

Continua...

Tabela 3. Predição dos valores das médias dos genótipos ordenadas pela média a posteriori e seus ranques pela metodologia Bayesiana e análises frequentistas com predição dos valores genotípicos (utilizando BLUPS), média ajustada e média aritmética para 55 linhagens de soja, considerando produtividade de grãos kg.ha⁻¹. Jaboticabal, SP, 2016.

Ranque da média a posteriori	Genótipos	Média a posteriori	⁽¹⁾ IC 95%		Ranque Médio da posteriori	Ranque a posteriori (Quantis)					⁽²⁾ Valor genotípico (u+g+gem)	Média fenotípica (<i>lsmeans</i>)	Média aritmética	⁽³⁾ N
			Inferi or	Super ior		0,025	0,05	0,5	0,95	0,975				
18	Jab 11	2439,03	2100	2855	22,78	4	5	22	44	46,98	2448,51	2510,01	2437,05	17
19	Jab 47	2429,59	2020	2778	23,14	4	6	22	44	47,98	2454,61	2482,74	2574,31	16
20	Jab 3	2420,53	2055	2788	23,55	4	6	23	44	47,98	2436,79	2481,02	2481,02	18
21	Jab 10	2419,02	2033	2821	23,87	4	5	23	45,98	49	2419,94	2470,63	2428,62	17
22	Jab 49	2410,54	2050	2807	24,20	4	6	23	45	48	2417,66	2450,87	2450,87	18
23	Jab 36	2398,55	2016	2772	25,00	5	6,05	24,25	45	49,98	2416,74	2420,38	2464,28	17
24	Jab 23	2391,44	1989	2749	25,68	4,03	6	25	47	49	2404,65	2430,35	2430,35	18
25	Jab 38	2383,61	2007	2750	26,05	4	6	25,25	47	49	2392,75	2411,59	2411,59	18
26	Jab 19	2358,46	1947	2712	27,54	5	7	27	49	50	2363,68	2364,64	2439,82	17
27	Jab 4	2357,72	1976	2728	27,73	5	7	28	47,48	49	2352,26	2313,35	2433,90	17
28	Jab 46	2355,90	1957	2718	27,78	5	7	27	48	50	2344,65	2327,44	2416,70	17
29	Jab 29	2355,75	2008	2763	27,94	5	7	29	47	49	2356,50	2341,90	2401,11	17
30	Jab 34	2351,34	1960	2728	28,22	5	7	29	48	50	2341,81	2279,17	2263,80	16
31	Jab 18	2340,73	1912	2671	28,57	6	9	28	49	50	2344,21	2335,06	2335,06	18
32	Jab 51	2329,56	1997	2717	29,52	7	9	30	48	50	2337,37	2324,28	2324,28	18
33	Jab 40	2303,98	1889	2657	30,97	6,03	10	32	50	51,49	2296,39	2259,66	2259,66	18
34	Jab 50	2294,55	1943	2680	31,76	7	11	33	49	51	2267,20	2206,48	2255,05	17

Continua...

Tabela 3. Predição dos valores das médias dos genótipos ordenadas pela média a posteriori e seus ranques pela metodologia Bayesiana e análises frequentistas com predição dos valores genotípicos (utilizando BLUPS), média ajustada e média aritmética para 55 linhagens de soja, considerando produtividade de grãos kg.ha⁻¹. Jaboticabal, SP, 2016.

Ranque da média a posteriori	Genótipos	Média a posteriori	⁽¹⁾ IC 95%		Ranque Médio da posteriori	Ranque a posteriori (Quantis)					⁽²⁾ Valor genotípico (u+g+gem)	Média fenotípica (<i>lsmeans</i>)	Média aritmética	⁽³⁾ N
			Inferi or	Super ior		0,025	0,05	0,5	0,95	0,975				
35	Jab 28	2293,54	1927	2659	31,54	9	11,5	32	50	52	2287,64	2245,86	2245,86	18
36	Jab 33	2277,31	1925	2683	32,53	7,01	11	33	51	52	2258,88	2200,53	2200,53	18
37	Jab 9	2276,19	1868	2670	32,82	7	10	34	51	52	2250,81	2182,33	2191,63	16
38	Jab 20	2276,16	1876	2629	32,76	8	12	34	51	52	2263,70	2227,69	2221,25	17
39	Jab 21	2253,82	1854	2629	34,10	9	13	35	51	53	2231,06	2156,66	2156,66	18
40	Jab 22	2233,35	1859	2620	35,37	10	13,05	37	52	53	2214,37	2130,35	2130,35	18
41	CD-216	2214,19	1846	2596	36,46	11	15,53	38	52	53	2163,76	2027,43	2063,77	16
42	Jab 8	2204,78	1842	2628	36,98	11,03	16	39	52	54	2175,53	2069,11	2069,11	18
43	Jab 13	2175,63	1798	2554	38,65	14,03	18	41	53	54	2160,64	2055,34	2108,44	16
44	Jab 16	2173,26	1770	2558	38,89	13	18	41	53	54	2140,16	2013,35	2013,35	18
45	V-Max	2148,40	1651	2722	38,56	7	11	42	55	55	2244,82	2234,26	2174,49	16
46	Jab 37	2139,46	1696	2492	40,57	15	20	42,75	54	55	2092,06	1905,08	1954,90	17
47	Jab 17	2112,46	1622	2591	40,85	12	16	44	55	55	2063,98	1103,34	1103,34	6
48	Jab 14	2107,45	1708	2476	42,34	18	24	44,75	54	55	2058,68	1884,88	1884,88	18
49	Jab 12	2105,07	1598	2541	41,19	13	17,53	44	55	55	2056,02	1614,92	1614,92	7
50	Jab 48	2098,04	1683	2485	42,53	19	23	45	54	55	2041,42	1857,68	1857,68	18
51	Jab 15	2037,59	1668	2451	45,45	23	28	48	55	55	1984,18	1767,43	1767,43	18

Continua...

Tabela 3. Predição dos valores das médias dos genótipos ordenadas pela média a posteriori e seus ranques pela metodologia Bayesiana e análises frequentistas com predição dos valores genotípicos (utilizando BLUPS), média ajustada e média aritmética para 55 linhagens de soja, considerando produtividade de grãos kg.ha⁻¹. Jaboticabal, SP, 2016.

Ranque da média a posteriori	Genótipos	Média a posteriori	⁽¹⁾ IC 95%		Ranque Médio da posteriori	Ranque a posteriori (Quantis)					⁽²⁾ Valor genotípico (u+g+gem)	Média fenotípica (<i>lsmeans</i>)	Média aritmética	⁽³⁾ N
			Inferi or	Super ior		0,025	0,05	0,5	0,95	0,975				
52	Jab 26	2015,36	1615	2368	46,09	25	29	49	55	55	1956,89	1724,40	1724,40	18
53	Jab 24	1969,07	1596	2377	48,08	29	33	50,5	55	55	1889,34	1617,90	1617,90	18
54	Jab 6	1938,13	1529	2351	48,83	29,03	34	51	55	55	1865,43	1586,84	1635,00	17
55	Jab 25	1908,55	1476	2283	49,74	33	37	52	55	55	1826,89	1472,25	1518,55	17

(1). Intervalo de credibilidade a posteriori, (2). Valor genotípico: média geral do experimento + BLUP + Interação média genótipo com ambiente, (3). Número de repetições total ao qual o genótipo foi avaliado.

Para os percentis 2,5 e 5% os genótipos Jab 48, Jab 15, Jab 26, Jab 24, Jab 6 e Jab 25 não atingiram uma alta classificação de ranque, ou seja, próximo de um, e não devem ser selecionados para cultivo. De acordo com os ranques para os percentis 95 ou 97,5% que mostram o quão improdutivo determinado genótipo seria caso fosse selecionando com base na média, o genótipo Jab 42 poderia descer para 2º (percentil 97,5%), ou se mantendo na primeira posição (percentil 95% e demais).

Para alguns genótipos existem discrepância entre as magnitudes dos valores de média para a abordagem Bayesiana, *Ismeans* e médias genotípicas quando comparados com a média aritmética (Tabela 3). Três genótipos apresentaram grandes diferenças entre as médias para os quatro métodos avaliados, aos quais são respectivamente: Jab 17: genótipo com alta perda de parcelas (N=6 parcelas remanescentes), Jab 25: genótipo com pior desempenho de produtividade de grãos (1908,55 kg.ha⁻¹) e Jab 42: genótipo com melhor desempenho (3331,33 kg.ha⁻¹) (Tabela 3). Portanto, em situações mais extremas existem diferenças entre os quatro métodos avaliados, ressaltando que a média aritmética é a que mais apresenta maior variação da produtividade de grãos dos genótipos.

4.3 Correlação entre as metodologias Bayesianas, modelos mistos via REML/BLUP, médias ajustadas (*Ismeans*) e médias aritméticas simples

As médias calculadas pelos métodos Bayesianos, *Ismeans* e médias genotípicas foram as menos discrepantes em termos de magnitudes e ranque, porque são métodos que levam em conta o delineamento adotado, fazendo ajuste nas médias quando necessário (Tabela 4). No método Bayesiano uma amostra da distribuição a posteriori é gerada para os parâmetros estudados (genótipos, herdabilidade, dentre outros) que são considerados como de efeito aleatório. No método *Ismeans* (Least-squares means) ou médias marginais preditas, os genótipos são tratados como sendo de efeito fixo, e o ajuste das médias é feito com base nos fatores presentes no modelo linear (genótipos, repetições/ambientes e ambientes). No método das médias genotípicas os genótipos são tratados como aleatórios, e a média de determinado genótipo para essa metodologia considerando experimentos multiambientes é igual a média geral do experimento + BLUP do genótipo + interação genótipos por ambientes média.

Tabela 4. Correlação entre as médias e ranques dos métodos avaliados. Diagonal superior, correlação de Pearson entre médias de genótipos e diagonal inferior, correlação de Spearman entre ranques dos genótipos, considerando produtividade de grãos kg.ha⁻¹. Jaboticabal, SP, 2016.

	Bayesiana	BLUPS	<i>Lsmeans</i>	Aritmética
Bayesiana	-	0,99742 <0,0001*	0,96775 <0,0001	0,96627 <0,0001
BLUPS	0,99740 <0,0001	-	0,97256 <0,0001	0,96945 <0,0001
<i>Lsmeans</i>	0,98954 <0,0001	0,99430 <0,0001	-	0,99629 <0,0001
Aritmética	0,98867 <0,0001	0,99091 <0,0001	0,99019 <0,0001	-

*: significância avaliada pelo teste t

Em termos de correlação entre as médias, os métodos Bayesianos e BLUPS foram similares, em que estes apresentaram $r=0,9974$. O método *Lsmeans* para cálculo de médias de genótipos apresentou a segunda maior correlação com o método Bayesiano, $r=0,9678$, e o método da média aritmética simples apresentou correlação $r=0,9663$ (Tabela 4). A metodologia REML/BLUP é considerada como o método padrão para cálculos de efeitos aleatórios no melhoramento de plantas na seleção de genótipos (OLIVEIRA et al., 2016). Em casos de experimentos levemente ou parcialmente balanceados este método pode apresentar alta correlação com o método Bayesiano.

No presente trabalho ocorreram 5% de parcelas perdidas e, portanto, existe desbalanceamento. Contudo, para efeitos de ranqueamento de genótipos as parcelas perdidas não influenciaram de forma significativa. Assim, é necessário avaliar o método Bayesiano em outras condições como grande de número de

parcelas perdidas e variâncias dos tratamentos heterogêneas para verificar sua eficácia.

4.5 Parâmetros genéticos para os métodos Bayesiano e modelos mistos

Na Tabela 5, as estimativas dos parâmetros genéticos para as abordagens Bayesianas e modelos mistos foram similares para herdabilidade em nível de média (h^2_{mg}), Acurácia na seleção dos genótipos (Acgen), Coeficiente de variação genético (CVgi%) e Coeficiente de variação ambiental (CVe%), além disso, os valores das herdabilidades e correlação genotípica entre o desempenho dos genótipos nos vários ambientes para a produtividade de grãos em $kg.ha^{-1}$ está bem próximo de valores relatado na literatura (GOMEZ, 2014). O valor da acurácia que foi acima de 0,7 para ambas as metodologias é classificada como alta (RESENDE; DUARTE, 2007).

Tabela 5. Estimativa de parâmetros genéticos para os métodos da verossimilhança e Bayesiano, considerando produtividade de grãos $kg.ha^{-1}$. Jaboticabal, SP, 2016.

Parâmetro Genético	REML	Bayesiano	IC 95% [inf.; Sup.]
σ_g^2	102.260,30	84.841,57	
σ_{ge}^2	142.192,01	4.433,44	
σ_e^2	1.312.588,37	1.269.348,00	
σ_f^2	1.557.040,68	1.354.189,60	
h^2_{mg}	0,51	0,53	(0,39;0,68)
Acgen	0,72	0,73	
c^2_{int}	0,09	0,003	
rgloc	0,42	0,95	
CVgi%	13,55	12,36	
CVe%	48,55	47,79	
PEV	49.680,15	18.544,99	
SEP	222,89	136,18	
$\bar{M}$	2360,06	2357,51	

$\sigma_f^2 = \sigma_g^2 + \sigma_{ge}^2 + \sigma_e^2$: Variância Fenotípica; σ_g^2 : Variância Genotípica; σ_{ge}^2 : Variância da interação G x E; $c^2_{int} = \sigma_{ge}^2 / \sigma_f^2$: Coeficiente de determinação dos efeitos da interação G x E; σ_e^2 : Variância Residual; $h^2_{mg} = (\sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2 / N.amb + \sigma_e^2 / (N.amb \times N.rep)))$: Herdabilidade em nível de média (genótipos); $Acgen = \sqrt{(1 - PEV / \sigma_g^2)}$: Acurácia na seleção dos genótipos; $CVgi = \sqrt{\sigma_g^2 / \bar{M}}$: Coeficiente de variação genético; $CVe = \sqrt{\sigma_e^2 / \bar{M}}$: Coeficiente de variação ambiental; $rgloc = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2)$: Correlação genotípica entre o desempenho dos genótipos nos vários ambientes; $PEV = (1 - Acgen^2) \times \sigma_g^2$: Erro de predição da variância; $SEP = \sqrt{PEV}$: Desvio padrão do valor genotípico da progênie; $\bar{M}$: Média Geral e IC: Intervalo de credibilidade. N.amb: Número de ambientes, N.rep: Número de repetições.

O método Bayesiano apresentou menor variância de interação genótipos x ambientes, produzindo menor valor do coeficiente de determinação dos efeitos da interação G x E (c^2_{int}). Consequentemente, aumentou-se a correlação genotípica entre o desempenho dos genótipos nos vários ambientes (rgloc), aproximando o valor da unidade. A implicação do rgloc próximo de um para seleção é do aumento da confiabilidade na seleção de determinado genótipo testado nos diversos ambientes avaliados.

A partir de estudos de simulação para interação G x E ficou definido que $rgloc \geq 0,80$ indica interação simples e $rgloc \leq 0,20$ indica complexa (CRUZ; CASTOLDI, 1991). Assim, quanto mais os valores de rgloc são próximos de 1, menos complexo um estudo de interação genótipos x ambientes e adaptabilidade e estabilidade fica, porque as variações ambientais são diminuídas (ROSADO et al., 2012). A existência da interação está associada a dois fatores: o primeiro, denominado simples, é proporcionado pela diferença de variabilidade entre genótipos nos ambientes; e o segundo, denominado complexo, indica a inconsistência da superioridade de genótipos com a variação ambiental, ou seja, haverá genótipos com desempenho superior em um ambiente, mas não em outro, tornando mais difícil atingir os objetivos de seleção (CRUZ; REGAZZI; CARNEIRO, 2012).

No método Bayesiano, o erro de predição das variâncias dos valores genéticos (PEV) e desvio padrão da variância dos valores genéticos (SEP) foram menores que no método de modelos mistos, indicando um melhor ajuste no cálculo das médias a posteriori dos genótipos de soja. Valores mais baixos de PEV e SEP

são desejáveis no melhoramento de plantas e estão diretamente relacionados com a precisão e maximização da acurácia (RESENDE; DUARTE, 2007).

Os parâmetros genéticos podem ser calculados a partir de diferentes tipos de análise de dados como ANOVA, REML (modelos mistos) e métodos Bayesianos. Para avaliação dos parâmetros calculados no método Bayesiano é feito o cálculo do intervalo de credibilidade. Este foi feito apenas para a herdabilidade média de genótipos. Contudo, pode ser calculada para qualquer estimativa de parâmetro genético da Tabela 5. No método REML é possível calcular um intervalo de confiança para herdabilidade utilizando um método denominado de delta (OEHLERT, 1992). Para intervalo de confiança e os outros parâmetros no REML outra alternativa pode ser a utilização de *bootstrapping*.

Assim, as vantagens dos métodos Bayesianos com dados de experimentação agrícolas é que podem permitir o uso de informações a priori obtidas de dados de experimentos anteriores, apresentam modelagem menos sensível a *outliers* do que os métodos frequentistas e não necessitam de muitas pressuposições para realização das análises (ex.: amostra grande, balanceamento, etc). Ressalta-se ainda que os métodos Bayesianos com dados de experimentação agrícola em experimentos complexos com vários anos, safras, locais, interações e aninhamentos apresentam um tempo de computação levemente maior que modelos mistos. Além disso, em casos específicos, com uso de priores informativas, as análises Bayesianas podem produzir distribuições à posteriori que são fortemente influenciados pelas distribuições a priori, devendo, portanto, serem utilizadas com parcimônia.

5 CONCLUSÕES

1) Os valores genotípicos obtidos pela abordagem Bayesiana, diferiram das médias com a inferência frequentista (modelos mistos, lsmeans e médias aritméticas simples), em experimentos multiambientais na cultura da soja para os genótipos com alta e baixa produtividade de grãos. Para os demais casos os quatro métodos testados foram equivalentes.

- 2) A abordagem Bayesiana com informações a priori de experimentos anteriores com soja pode ser utilizada como uma ferramenta no melhoramento genético dessa cultura.
- 3) O método de modelos mistos (REML) apresentou valores de médias e parâmetros genéticos próximos do método Bayesiano para herdabilidade em nível de média (h^2_{mg}), Acurácia na seleção dos genótipos (Ac_{gen}), Coeficiente de variação genético ($CV_{gi}\%$) e Coeficiente de variação ambiental ($CV_e\%$);
- 4) É necessário avaliar o método Bayesiano em outras condições como: grande número de parcelas perdidas e variâncias dos tratamentos heterogêneas.

REFERÊNCIAS

ALMEIDA, G. C. S.; LAMOUNIER, W. M. Os alimentos transgênicos na agricultura brasileira: evolução e perspectivas. **Organizações Rurais & Agroindustriais**, Lavras, v. 7, n. 3, p. 345-355, 2011.

ALMEIDA, L. A.; KIIHL, R. A. D. S.; MIRANDA, M. A. C.; CAMPELO, G. J. D. A. Melhoramento da soja para regiões de baixas latitudes. In: QUEIRÓZ, M. A.; GOEDERT, C. O.; RAMOS, S. R. R. **Recursos Genéticos e Melhoramento de Plantas para o Nordeste Brasileiro**. Petrolina: Embrapa Semii-Árido, v.1, 1999. p. (online). Disponível em: < <http://www.cpatsa.embrapa.br/catalogo/livrorg/> >. Acesso em: 29 dez. 2014.

ANDRADE, A. C. B.; DA SILVA, A. J.; FERRAUDO, A. S.; UNÊDA-TREVISOLI, S. H.; DI MAURO, A. O. Strategies for selecting soybean genotypes using mixed models and multivariate approach. **African Journal of Agricultural Research**, v. 11, n. 1, p. 23-31, 2016. Disponível em: < <http://www.academicjournals.org/journal/AJAR/article-full-text-pdf/CD2848756664> >. Acesso em: 05 mar. 2017.

BACAXIXI, P.; RODRIGUES, L. R.; BRASIL, E. P.; BUENO, C. E. M. S.; RICARDO, H. A.; EPIPHANIO, P. D.; SILVA, D. P.; BARROS, B. M. C.; SILVA, T. F.; BOSQUÊ, G. G. A soja e seu desenvolvimento no melhoramento genético. **Revista Científica Eletrônica de Agronomia**, Garça, v. 10, n. 20, p. 1-6, 2011.

BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. Fitting linear mixed-effects models using lme4. **Journal of Statistical Software**, v. 67, n. 1, p. 1-48, 2014.

BORÉM, A.; CONDORI, M.; MIRANDA, G. V. **Mejoramiento de Plantas**. 4.ed. Viçosa, MG: Editora UFV, 2008. 438 p.

BORGES, V.; FERREIRA, P. V.; SOARES, L.; SANTOS, G. M.; SANTOS, A. M. M. Seleção de clones de batata-doce pelo procedimento REML/BLUP. **Acta Scientiarum. Agronomy**, Maringá, v. 32, n. 4, p. 643-649, 2010.

CARNEIRO JÚNIOR, J. M.; DE ASSIS, G. M. L.; EUCLYDES, R. F.; LOPES, P. S. Influência da informação a priori na avaliação genética animal utilizando dados simulados. **Revista Brasileira de Zootecnia**, v. 34, n. 6, p. 1905-1913, 2005.

CARNEIRO JÚNIOR, J. M.; DE ASSIS, G. M. L.; EUCLYDES, R. F.; MARTINS, W. M. D. O.; WOLTER, P. F. Predição de valores genéticos utilizando inferência bayesiana e frequentista em dados simulados. **Acta Scientiarum. Animal Sciences**, v. 32, n. 3, p. 337-344, 2010.

CARVALHO, A. D. F.; FRITSCH NETO, R.; GERALDI, I. O. Estimation and prediction of parameters and breeding values in soybean using REML/BLUP and Least Squares. **Crop Breeding and Applied Biotechnology**, v. 8, n. 3, p. 219-224, 2008.

COMPANHIA NACIONAL DE ABASTECIMENTO. Séries Históricas de Área Plantada, Produtividade e Produção. **Companhia Nacional de Abastecimento**, 2017. Disponível em: <
http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17_02_16_11_51_51_boletim_graos_fevereiro_2017.pdf>. Acesso em: 05 mar. 2017.

COSTA, M. T. G. P.; SANCHES, A.; MUNARI, D. P. Estimacão bayesiana de parâmetros genéticos de pesos corporais em um rebanho da raça Guzerá. **Nucleus Animalium**, v. 1, n. 1, p. 1-13, 2009.

CRUZ, C. D.; CASTOLDI, F. L. Decomposição da interação genótipos x ambientes em partes simples e complexa. **Revista Ceres**, v. 38, n. 219, 1991. Disponível em: <
<http://www.ceres.ufv.br/ojs/index.php/ceres/article/view/2165/203>>. Acesso em: 12 out. 2016.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos Biométricos Aplicados ao Melhoramento Genético**. vol. 1. 4. ed. Viçosa, MG: UFV, 2012. 514 p.

DUARTE, J. B.; ZIMMERMANN, M. J. D. O. Correlation among yield stability parameters in common bean. **Crop Science**, v. 35, n. 3, p. 905-912, 1995.

FREITAS, M. D. C. M. D. A cultura da soja no Brasil: o crescimento da produção brasileira e o surgimento de uma nova fronteira agrícola. **Enciclopédia Biosfera, Centro Científico Conhecer: Goiânia**, v. 7, n. 12, p. 1-12, 2011.

GELMAN, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). **Bayesian analysis**, v. 1, n. 3, p. 515-534, 2006.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical science**, p. 457-472, 1992.

GODOI, C. R. C. D. **Divergência genética e predição de valores genotípicos em soja**. 2014. 108 f. Tese (Doutorado em Genética e Melhoramento de Plantas)-Universidade Federal de Goiás, Escola de Agronomia, Goiânia, GO, 2014.

GOMEZ, G. M. **Abordagem genética e multivariada na performance agronômica de genótipos de soja oriundos de diferentes genealogias**. 2014. 79 f. Tese (Doutorado)-Universidade Estadual Paulista "Júlio de Mesquita Filho", Produção Vegetal, Jaboticabal, SP, 2014.

GONÇALVES, M. C.; FRITSCH NETO, R. **Tópicos especiais de biometria no melhoramento de plantas**. 1ª Ed. Visconde do Rio Branco: Suprema, 2012. 282 p.

LAM, H.-M.; XU, X.; LIU, X.; CHEN, W.; YANG, G.; WONG, F.-L.; LI, M.-W.; HE, W.; QIN, N.; WANG, B. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. **Nature genetics**, v. 42, n. 12, p. 1053-1059, 2010.

MAIA, M. C. C.; VELLO, N. A.; ARAÚJO, L. B. D.; DIAS, C. T. D. S.; OLIVEIRA, L. C. D.; ROCHA, M. D. M. Interação genótipo x ambiente com uso da análise de componentes principais para populações de soja selecionadas para resistência a insetos. **Revista Brasileira de Biometria**, v. 31, n. 1, p. 13-27, 2013.

NOGUEIRA, D. A.; SÁFADI, T.; BEARZOTI, E.; BUENO-FILHO, J. Análises clássica e Bayesiana de um modelo misto aplicado ao melhoramento animal: uma ilustração. **Ciência e Agrotecnologia, Lavras**, v. 27, p. 1614-1624, 2003.

NUNES, G. H. S. I. **Interação genótipos x ambientes em eucalipto: implicações sobre a seleção e formas de atenuar seu efeito**. 2000. 160 f. Tese (Doutorado)-Universidade Federal de Lavras, Lavras, MG, 2000.

OEHLERT, G. W. A note on the delta method. **The American Statistician**, v. 46, n. 1, p. 27-29, 1992.

OLIVEIRA, G. H. F.; BUZINARO, R.; REVOLTI, L. T. M.; GIORGENON, C. H. B.; CHARNAL, K.; RESENDE, D.; MORO, G. V. An accurate prediction of maize crosses using diallel analysis and best linear unbiased predictor (BLUP). **Chilean Journal of Agricultural Research**, v. 76, n. 3, 2016.

OLIVEIRA, H. N. D.; LÔBO, R. B.; PEREIRA, C. S. Comparação de modelos não-lineares para descrever o crescimento de fêmeas da raça guzerá. **Pesquisa Agropecuária Brasileira**, v. 35, n. 9, p. 1843-1851, 2000.

OMER, S. O.; ABDALLA, A. H.; CECCARELLI, S.; GRANDO, S.; SINGH, M. Bayesian estimation of heritability and genetic gain for subsets of genotypes evaluated in a larger set of genotypes in a block design. **European Journal of Experimental Biology**, v. 4, n. 3, p. 566-575, 2014a.

OMER, S. O.; ABDALLA, A. W. H.; MOHAMMED, M. H.; SINGH, M. Bayesian estimation of genotype-by-environment interaction in sorghum variety trials. **Communications in Biometry and Crop Science**, v. 10, p. 82-95, 2015.

OMER, S. O.; ABDALLA, A. W. H.; SARKER, A.; SINGH, M. Bayesian determination of the number of replications in crop trials. **European Journal of Experimental Biology**, v. 4, n. 6, p. 129-133, 2014b.

PACHECO, R. M.; DUARTE, J. B.; DA SILVA ASSUNÇÃO, M.; JÚNIOR, J. N. Zoneamento e adaptação produtiva de genótipos de soja de ciclo médio de maturação para Goiás. **Pesquisa Agropecuária Tropical (Agricultural Research in the Tropics)**, v. 33, n. 1, p. 23-27, 2007.

PAIVA, B. M. D.; ALVES, R. M.; HELENO, N. M. Aspecto socioeconômico da soja. **Informe agropecuário**, Belo Horizonte, v. 27, n. 230, p. 7-14, 2006.

PELUZIO, J. M.; DE DEUS GEROMINNI, G.; DA SILVA, J. P. A.; AFFÉRRRI, F. S.; VENDRUSCOLO, J. B. G. Estratificação e dissimilaridade ambiental para avaliação de cultivares de soja no estado de Tocantins. **Bioscience Journal**, v. 28, n. 3, 2012.

PEREIRA, É. M.; GRAVINA, G. A.; THIÉBAUT, J. T. L. Aplicações dos modelos lineares mistos na pesquisa agropecuária. **Natureza on line**, v. 10, n. 2, p. 52-58, 2012.

PINHEIRO, L. C. D. M.; GOD, P. I. V. G.; FARIA, V. R.; OLIVEIRA, A. G.; HASUI, A. A.; PINTO, E. H. G.; ARRUDA, K. M. A.; PIOVESAN, N. D.; MOREIRA, M. A. Parentesco na seleção para produtividade e teores de óleo e proteína em soja via modelos mistos. **Pesquisa Agropecuária Brasileira**, v. 48, p. 1246-1253, 2013. Disponível em: < http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2013000900008&nrm=iso >. Acesso em: 05 mar. 2017.

PRADO, E. E. D.; HIROMOTO, D. M.; GODINHO, V. D. P. C.; UTUMI, M. M.; RAMALHO, A. R. Adaptabilidade e estabilidade de cultivares de soja em cinco épocas de plantio no cerrado de Rondônia. **Pesquisa Agropecuária Brasileira**, v. 36, n. 4, p. 625-635, 2001.

PRIOLLI, R. H. G.; MENDES-JUNIOR, C. T.; SOUSA, S. M. B.; SOUSA, N. E. A.; CONTEL, E. P. B. Diversidade genética da soja entre períodos e entre programas de melhoramento no Brasil. **Pesquisa Agropecuária Brasileira**, v. 39, n. 10, p. 967-975, 2004.

PRIOLLI, R. H. G.; WYSMIERSKI, P. T.; DA CUNHA, C. P.; PINHEIRO, J. B.; VELLO, N. A. Genetic structure and a selected core set of Brazilian soybean cultivars. **Genetics and Molecular Biology**, v. 36, n. 3, p. 382-390, 2013. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3795172/> >. Acesso em: 12 out. 2016.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, 2016. Disponível em: < <http://www.R-project.org/> >. Acesso em: 17 out. 2016.

RAMALHO, M. P.; SANTOS, J. B.; ZIMMERMANN, M. J. Interação dos genótipos por ambientes. In: RAMALHO, M. P.; SANTOS, J. B.; ZIMMERMANN, M. J. **Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro**. Goiânia: Editora da UFG, 1993. p. 137-170.

RESENDE, M. D. V. **Métodos estatísticos ótimos na análise de experimentos de campo**. Colombo: Embrapa Florestas, 2004. 65 p. (Embrapa Florestas. Documentos, 100).

RESENDE, M. D. V. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas, 2007a. 561 p.

RESENDE, M. D. V. **SELEGEN-REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos**. Colombo: Embrapa Florestas, 2007b. 359 p. Disponível em: < http://www.incaper.es.gov.br/congressos/cbmp/apresentacoes/minicursos/Minicurso2_SelegenManual.pdf >. Acesso em: 29 set. 2016.

RESENDE, M. D. V. D.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v. 37, n. 3, p. 182-194, 2007.

ROSADO, A. M.; ROSADO, T. B.; ALVES, A. A.; LAVIOLA, B. G.; BHERING, L. L. Seleção simultânea de clones de eucalipto de acordo com produtividade, estabilidade e adaptabilidade. **Pesquisa Agropecuária Brasileira**, v. 47, n. 7, p. 964-971, 2012.

ROSSMANN, H. **Estimativas de parâmetros genéticos e fenotípicos de uma população de soja avaliada em quatro anos**. 2001. 80 f. Dissertação (Mestrado)-Escola Superior de Agricultura Luiz de Queiroz, Genética e Melhoramento, Piracicaba, 2001.

SHEN, Y.; LIU, X. Phenological Changes of Corn and Soybeans over US by Bayesian Change-Point Model. **Sustainability**, v. 7, n. 6, p. 6781-6803, 2015.

SILVA, F. F. E.; VIANA, J. M. S.; FARIA, V. R.; RESENDE, M. D. V. D. Bayesian inference of mixed models in quantitative genetics of crop species. **Theoretical and applied genetics**, v. 126, n. 7, p. 1749-1761, 2013a.

SILVA, M.; PETERNELLI, L.; NASCIMENTO, M.; DA SILVA, F. L. Modelos mistos na seleção de famílias de cana-de-açúcar aparentadas sob o enfoque clássico e bayesiano. **Revista Brasileira de Biometria**, v. 31, p. 1-12, 2013b.

SILVEIRA, G. D.; DI MAURO, A. O.; CENTURION, M. A. P. D. C. Seleção de genótipos de soja para a região de Jaboticabal (SP)(Ano agrícola 23-24). **Científica**, Jaboticabal, v. 34, n. 1, p. 92-98, 2008.

SINGH, M.; AL-YASSIN, A.; OMER, S. O. Bayesian Estimation of Genotypes Means, Precision, and Genetic Gain Due to Selection from Routinely Used Barley Trials. **Crop Science**, v. 55, n. 2, p. 501-513, 2015.

SLEPER, D. A.; SHANNON, J. G. Role of public and private soybean breeding programs in the development of soybean varieties using biotechnology. **AgBioForum**, v. 6, p. 27-32, 2003.

SORENSEN, D.; GIANOLA, D. **Likelihood, bayesian and MCMC methods in quantitative genetics**. Amsterdam: Springer, 2002. 740 p.

STURTZ, S.; LIGGES, U.; GELMAN, A. R2WinBUGS: a package for running WinBUGS from R. **Journal of Statistical software**, v. 12, n. 3, p. 1-16, 2005.

VALENTIM, F. L.; SILVA, R. M. D. A.; ALVES, M. D. C. Modelos de redes bayesiana para incidência da ferrugem asiática na soja, cultivar suprema, em diferentes condições de temperatura e molhamento foliar. In: XXXIX SBPO - A Pesquisa Operacional e o Desenvolvimento Sustentável, 2007, Fortaleza, CE, Brasil, p. 2677-2678.

ZEILEIS, A.; HOTHORN, T. Diagnostic Checking in Regression Relationships. **R News**, v. 2, n. 3, p. 7-10, 2002. Disponível em: < <http://CRAN.R-project.org/doc/Rnews/> >. Acesso em: 19 out. 2016.