

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

**INTERAÇÃO GENÓTIPO X AMBIENTE, ADAPTABILIDADE E
ESTABILIDADE EM GENÓTIPOS DE CANA-DE-AÇÚCAR**

Flávio Cese Arantes
Engenheiro Agrônomo

2013

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

**INTERAÇÃO GENÓTIPO X AMBIENTE, ADAPTABILIDADE E
ESTABILIDADE EM GENÓTIPOS DE CANA-DE-AÇÚCAR**

Flávio Cese Arantes

Orientador: Prof. Dr. Dilermando Perecin

Coorientador: Dr. Marcos Guimarães de Andrade Landell

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

A662i Arantes, Flávio Cese
Interação genótipo x ambiente, adaptabilidade e estabilidade em
genótipos de cana-de-açúcar / Flávio Cese Arantes. – – Jaboticabal,
2013
xii, 94 f. ; 28 cm

Tese - Universidade Estadual Paulista, Faculdade de Ciências
Agrárias e Veterinárias, 2013
Orientador: Dilermando Perecin
Coorientador: Marcos Guimarães de Andrade Landell
Banca examinadora: Euclides Braga Malheiros, Rinaldo Cesar de
Paula, Maximiliano Salles Scarpari, Ivana Marino Bárbaro
Bibliografia

1. Modelos Mistos 2. Bayesiana 3. MHPRVG I. Título. II.
Jaboticabal - Faculdade de Ciências Agrárias e Veterinárias.

CDU 631.52:633.61

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

CERTIFICADO DE APROVAÇÃO

TÍTULO: INTERAÇÃO GENÓTIPO X AMBIENTE, ADAPTABILIDADE E ESTABILIDADE EM GENÓTIPOS DE CANA-DE-AÇÚCAR

AUTOR: FLÁVIO CESE ARANTES

ORIENTADOR: Prof. Dr. DILERMANDO PERECIN

CO-ORIENTADOR: Prof. Dr. MARCOS GUIMARÃES DE ANDRADE LANDELL

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


Prof. Dr. DILERMANDO PERECIN

Departamento de Ciências Exatas / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal


Prof. Dr. EUCLIDES BRAGA MALHEIROS

Departamento de Ciências Exatas / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal


Prof. Dr. RINALDO CÉSAR DE PAULA

Departamento de Produção Vegetal / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal


Prof. Dr. MAXIMILIANO SALLES SCARPARI

Instituto Agronômico de Campinas / Ribeirão Preto/SP


Profa. Dra. IVANA MARINO BARBARO

Agência Paulista de Tecnologia dos Agronegócios / Colina/SP

Data da realização: 06 de dezembro de 2013.

DADOS CURRICULARES DO AUTOR

Flávio Cese Arantes, nascido em Pereira Barreto, Estado de São Paulo, Brasil, no dia 28 de janeiro de 1984. Tornou-se Engenheiro Agrônomo pela Universidade Estadual Paulista (UNESP), Câmpus de Ilha Solteira, em 2008, onde também tornou-se Mestre na área de Agronomia (Genética e Melhoramento de Plantas) em 2010. Realizou, sob financiamento da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, parte do seu Doutorado na Universidade de Wisconsin, Estados Unidos da América, durante Janeiro a Agosto de 2012, em Novembro de 2013 foi contratado pela Bayer S/A como Coordenador de Pesquisa e em Dezembro de 2013 tornou-se Doutor em Agronomia (Genética e Melhoramento de Plantas) pela UNESP, Câmpus de Jaboticabal.

Dedico esta obra aos meus Avós

José Cese (In Memoriam) e Aparecida Alves Cese

Aos meus Pais

Altino Arantes Junior e Genifer Aparecida Cese

e

Aos meus Irmãos

André e Fabiano

..:Numen Lumen:..

Agradecimentos

Agradeço a Deus, tudo e a Todos a cada pôr do sol. Porém gostaria de deixar tais agradecimentos por escrito nesta Tese.

Agradeço:

A Deus, por sempre iluminar o meu caminho.

*Ao meu querido Avô, **José Cese** (In Memoriam), que me ensinou o que é honestidade, amor e dedicação por ser o exemplo disso tudo. Minha querida **Vó Cida**, sempre dedicada aos netos.*

*Aos meus Pais, **Altino e Genifer**, por serem o princípio de tudo aquilo que eu sou e que me tornei, pela vida, pelos ensinamentos e por infinitos motivos indescritíveis.*

*Aos meus irmãos, **André e Fabiano**, por serem os parceiros da vida, mais do que amigos, o Elo de minha jornada.*

Ao Professor Dilermando Perecin pela serenidade, sabedoria na orientação e exemplo de pessoa e profissional para mim.

Ao Pesquisador Dr. Marcos Landell pela coorientação da realização deste trabalho.

Aos Professores Drs. do Depto. Ciências Exatas da UNESP/FCAV, em especial: Antônio S. Ferraudó, Danísio P. Munari e João A. de Oliveira.

Aos Professores(as) Doutores(as) membros da banca do exame de qualificação: Gustavo Môro, Luciana Pinto, Rinaldo de Paula e Sandra Unêda-Trevisoli.

Aos Professores(as) Doutores(as) membros da comissão examinadora da Tese: Euclides Malheiros, Ivana Bárbaro, Maximiliano Scarpari e Rinaldo de Paula.

Ao Professor Guilherme Rosa pela oportunidade de realização de parte deste trabalho nos EUA.

Aos amigos adquiridos ao longo dessa difícil jornada de Pós graduação: Guilherme Ferraudó, Karen Alvim, Mauro Xavier, Paolo Zancanaro, Rafael Finholdt e Sybelli Espindola.

Aos Amigos(as): Karen Zancanaro, Leo Seno, Mirelle, Carol, Kelly, Salvador, Guilherme, Val, Marcão, Natália, Camilitcha, Giovanna, Priscila, Tati, Diego, Helio e Deise Junqueira, João Flávio, Danilo.

À Minha querida amiga e companheira, pelo amor, carinho e serenidade: Bárbara Bonfim.

An special gratefulness for those friends found along the way abroad: Afshin Samia, Ana Ruth, Andrea and Rafael Zortea, Axel Madera, Bruno Valente, Chuanyu Sun, Diogo Garcia, Fabiane Silva, Fabio Pinaff, Fernanda Lopes, Francisco Peñagaricano (Punch'o), German Muttoni and Rosina, Giovanni Baez, Huihui Duan, Leane Oliveira, Leon, Llibertat Tussell, Marcos Lagrotta, Paulino Perez, Rocha, Rocio Gamba and Cris, Saleh Shahinfar, Stella Salvo, Vivian Felipe, Yao Chen and YaoDong.

E a Todos aqueles que contribuíram direta ou indiretamente para a construção do meu ser e da confecção desta Tese.

Agradeço-Os.

*"Tudo passa... Mas a Nobreza uma vez que adquirida ao longo
da vida torna-se intrínseca à Alma. "*

O Autor

SUMÁRIO

RESUMO.....	XI
ABSTRACT.....	XII
CAPÍTULO 1 - Considerações gerais	1
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	2
2.1. A cana-de-açúcar (<i>Saccharum</i> spp.)	2
2.2. Modelos Mistos: efeitos fixos e aleatórios, Máxima Verossimilhança Restrita (REML) e Melhor Predição Linear não-viesada (BLUP).....	3
2.2.1. Efeitos fixos e aleatórios.....	3
2.2.2. Estimacão de componentes de variância via Máxima Verossimilhança Restrita (REML)	4
2.2.3. Melhor Predição Linear não-viesada (BLUP)	7
2.3. Inferência Bayesiana de componentes de variância	12
2.3.1. Amostrador de Gibbs	14
2.4. Interaçao genótipo x ambiente, adaptabilidade e estabilidade	16
2.4.1. Regressão Linear Fenotípica Conjunta (EBERHART; RUSSELL, 1966)	18
2.4.2. Média Harmônica da Performance Relativa do Valor Genotípico (RESENDE, 2004)	21
2.4.3. Norma de reação com covariáveis desconhecidas (SU et al., 2006).....	23
3.REFERÊNCIAS BIBLIOGRÁFICAS	26
CAPÍTULO 2 - MATERIAL E MÉTODOS.....	36
1. MATERIAL	36
2. MÉTODOS.....	36
2.1. Regressão Linear Fenotípica Conjunta (EBERHART; RUSSELL, 1966).....	36
2.2. Média Harmônica da Performance Relativa do Valor Genotípico (RESENDE, 2004)	38
2.3. Norma de reação com covariáveis desconhecidas (SU et al., 2006)	43
3. REFERÊNCIAS BIBLIOGRÁFICAS	43
CAPÍTULO 3 - Reaction norm with unknown covariates and Eberhart-Russell models to sugarcane's multi-environment trial	44
ABSTRACT	44
1. INTRODUCTION	44
2. MATERIAL AND METHODS	46
2.1. Sugar cane data set	46
2.2. Linear reaction norm model with unknown covariates	47
2.3. Prior Specifications.....	47
2.4. Bayesian Inference.....	48
2.5. Variance components and Genetic parameters.....	48
2.6. Eberhart-Russell Regression Analysis.....	49
2.7. Reaction norm and Breeding	49
3. RESULTS AND DISCUSSION	50
4. CONCLUSIONS	62
5. ACKNOWLEDGMENTS	63
6. REFERENCES	63

CAPÍTULO 4 - Média harmônica da performance relativa do valor genotípico em ensaios de múltiplos ambientes na cultura da cana-de-açúcar	66
1. INTRODUÇÃO	66
2. MATERIAL E MÉTODOS	67
3. RESULTADOS E DISCUSSÕES.....	67
3.1. Estimativa de componentes de variância, de parâmetros genéticos e Média Harmônica da Performance Relativa do Valor Genotípico (MHPRVG) via BLUP	67
3.1.1. Estimativa de componentes de variância e de parâmetros genéticos	67
3.1.2. Seleção conjunta para produtividade, estabilidade e adaptabilidade via MHPRVG	69
4. CONCLUSÕES	75
5. REFERÊNCIAS BIBLIOGRÁFICAS	76
CAPÍTULO 5 - Considerações finais e conclusões.....	76
1. CONSIDERAÇÕES FINAIS	77
2. CONCLUSÕES	83
3. REFERÊNCIAS BIBLIOGRÁFICAS	84
APÊNDICE	85
1. Diagnósticos de convergência	86

INTERAÇÃO GENÓTIPO X AMBIENTE, ADAPTABILIDADE E ESTABILIDADE EM GENÓTIPOS DE CANA-DE-AÇÚCAR

RESUMO - A cana-de-açúcar (*Saccharum* spp.) é a cultura mais aplicada para a produção de açúcar tanto para consumo interno quanto para exportação, além da produção de etanol e biomassa. A área de cultivo vem crescendo expressivamente, fato que leva melhoristas a buscar genótipos mais produtivos às diferentes áreas de plantio. Experimentos em multi ambientes são comumente conduzidos em programas de melhoramento de plantas principalmente para avaliação de genótipos potenciais para produção. Este estudo incluiu 16 genótipos de cana-de-açúcar em 14 locais e avaliados por quatro anos, ou seja, quatro cortes (cana planta, primeira, segunda e terceira soca). O objetivo foi investigar a adaptabilidade e estabilidade dos genótipos de cana-de-açúcar para o atributo Tonelada de Pol por Hectare (TPH) em experimentos multi ambientais a partir de três metodologias estatísticas: Regressão linear fenotípica conjunta (E-R), Norma de reação (NR) e Média harmônica da performance relativa do valor genotípico (MHPRVG). Os valores da herdabilidade média genotípica estimados via MHPRVG apresentaram-se altos, o que leva a altas acurácias na seleção genotípica (superiores a 90%). Os genótipos IACSP963076, IACSP963048 e IACSP985024, primeiro corte, IACSP963048, IACSP963076, IACSP962073 e IACSP985024, segundo corte, IACSP962073, IACSP962008 e IACSP963076, terceiro corte e IACSP963076, IACSP985024, IACSP962073 e IACSP962008, quarto corte, foram os selecionados quando se considera os atributos de estabilidade, adaptabilidade e produtividade (TPH), simultaneamente e, portanto, são os genótipos recomendados para o plantio nesta macro região estudada quando analisadas via MHPRVG. Em termos de valores médios, os genótipos recomendados pela seleção simultânea baseada na MHPRVG são os genótipos IACSP963076, IACSP962073, IACSP985024 e IACSP962008, alcançando um valor médio de 2,8% acima da média geral dos quatro cortes. A porcentagem coincidente de ordenamento a partir da E-R e NR foi de 31, 44, 50 e 56% para o primeiro (cana planta), segundo, terceiro e quarto corte, respectivamente. Ambas E-R e NR assistem a alocação de genótipos nos ambientes avaliados para plantio e na estabilização de objetivos de melhoramento. A herdabilidade para o atributo TPH estimada via NR é dependente do gradiente ambiental e aumenta conforme os cortes. Todos os métodos analisados nesta tese obtiveram bons resultados e podem ser utilizados na avaliação da interação genótipo x ambiente (GxA), adaptabilidade e estabilidade; o método E-R pode ser utilizado a princípio para uma visão geral do comportamento dos genótipos nos ambientes visto que fornece estimativas de coeficientes de regressão de valores fenotípicos; os métodos E-R e NR proporcionam uma melhor visualização gráfica da interação GxA, adaptabilidade e estabilidade; os métodos NR e MHPRVG possibilitam o uso de dados desbalanceados, utilizam o fator genótipo como aleatório, o que proporciona a estimação do valor genotípico do indivíduo e aumento na acurácia de seleção genotípica; a estimativa dos Índices Ambientais (I_j) foram semelhantes para as metodologias de E-R e NR, para este conjunto de dados analisados; o método NR estima coeficientes de regressão que predizem valores genotípicos, o que permite aos melhoristas uma melhor exploração do valor genotípico do indivíduo no programa de melhoramento genético.

Palavras-chave: melhoramento genético de cana-de-açúcar, MHPRVG, Norma de reação, REML/BLUP

GENOTYPE X ENVIRONMENT INTERACTION, ADAPTABILITY AND STABILITY IN SUGARCANE GENOTYPES

ABSTRACT - The sugarcane (*Saccharum* spp.) is the most applied crop in the sugar yielding as much as internal consumption as exportation, besides its application in yielding alcohol and biomass production. The crop area has been increased expressively leading breeders to search more yielding genotypes to the different planted areas. Multi-environment trials are commonly conducted in plant breeding programs to evaluate potential genotypes mainly to yield. This study included 16 sugarcane genotypes at 14 locations evaluated in four years, that is, four harvests (sugarcane plant, first, second and third ratoon). The objective was to investigate the sugarcane's genotype adaptability and stability to Ton of Pol per Hectare (TPH) attribute in multi-environments trials from three statistical methodologies: Joint phenotypic linear regression (E-R), Reaction norm (NR) and Harmonic average performance relative breeding values (MHPRVG). The estimated genotypic average heritabilities values via MHPRVG showed themselves as high, what lead to high accuracy in the genotypic selection (higher than 90%). The genotypes IACSP963076, IACSP963048 and IACSP985024, first harvest, IACSP963048, IACSP963076, IACSP962073 and IACSP985024, second harvest, IACSP962073, IACSP962008 and IACSP963076, third harvest and IACSP963076, IACSP985024, IACSP962073 and IACSP962008, fourth harvest, were the selected when one considers simultaneously the stability, adaptability and productivity (TPH) attributes and thereby they are the recommended genotypes to be planted in this studied macro region when analyzed via MHPRVG. In a mean expression, the recommended genotypes by simultaneous selection based in MHPRVG are the IACSP963076, IACSP962073, IACSP985024 and IACSP962008 genotypes, reaching a average value of 2.8% higher than the four harvests overall mean. The percentage of selection ranking from both approach was 31, 44, 50 and 56% at the first (sugarcane plant), second, third and fourth harvest respectively. Both methodologies assist the allocation of the genotypes in the environments available for planting and to establish breeding objectives. In sugarcane, the TPH attribute heritability is dependent of environmental gradient and increase along the harvests. All the analyzed methods in this theses got good results e may be used in the genotypic x environment interaction (GxE) evaluation, adaptability and stability; the E-R method may be used at the beginning for the general view of the genotypes behavior along the environments once it provides estimations of the phenotypic slopes; both E-R and NR provide a better graphic view of the GxE interaction, adaptability and stability; the NR and MHPRVG methods provide the use of unbalanced data, assuming the genotypic factor as random, what leads the individual genotypic value estimation and genotypic selection accuracy increasing. The environmental index (I_j) estimation had a resemblance to E-R and NR methodologies to this analyzed data set; the NR method estimates slopes that predict the genotypic values allowing the breeders a better exploration of the individual genotypic value in a breeding program.

Keywords: MHPRVG, Reaction norm, REML/BLUP, Sugarcane breeding,

CAPÍTULO 1 - Considerações gerais

1. INTRODUÇÃO

A importância socioeconômica do cultivo da cana-de-açúcar (*Saccharum* spp.) é um fato observado desde os tempos coloniais da implantação desta cultura no Brasil. É a cultura mais utilizada na produção de açúcar para o consumo interno e produto de exportação, além de ser utilizada para a produção de álcool e biomassa. A cana-de-açúcar também tem demonstrado importância estratégica e ambiental, pela produção do biocombustível como fonte renovável de energia, menos poluente que combustíveis derivados do petróleo, e sua captação de carbono pela alta produção de biomassa. A área cultivada de cana-de-açúcar vem crescendo significativamente no Brasil, o que leva os melhoristas a buscarem genótipos que exibam maiores produtividades nas diferentes regiões de plantio.

A produção de cana-de-açúcar se concentra nas regiões Centro-Sul (RS, SC, PR, SP, RJ, MG, ES, MS, MT, GO e DF) e Norte-Nordeste do Brasil (BA, SE, AL, PE, PB, RN, CE, PI, MA, TO, PA, AP, RO, AM, AC e RR) e totaliza 588.478 mil toneladas de cana-de-açúcar na safra de 2012/2013. Segundo informações do Instituto Brasileiro de Geografia e Estatística (IBGE) a área plantada com a cultura da cana-de-açúcar no ano de 2011 foi de 9,6 milhões de hectares (UNICA, 2013).

Os programas de melhoramento genético visam suprir as necessidades crescentes desse mercado, a partir do desenvolvimento de novas variedades de cana-de-açúcar mais produtivas, com maior teor de sacarose e precocidade, resistentes a pragas, doenças e estresses abióticos (MATSUOKA et al., 2005).

Visto a grande extensão das áreas de produção de cana-de-açúcar e a demanda por genótipos mais produtivos e também a existência de fortes interações dos genótipos com os ambientes em que são plantados, quando se trata de características quantitativas como a produção, por exemplo, é importante, em programas de melhoramento, que se utilize métodos que explorem e capitalizem a interação genótipo por ambiente (GxA) para maximizar a produção.

A resposta diferenciada dos genótipos nos vários ambientes, conhecida como interação de genótipos com ambientes – GxA (EBERHART; RUSSELL, 1966) é um fenômeno natural que faz parte da evolução das espécies. Seus efeitos permitem o

aparecimento de genótipos estáveis e aptos a ambientes específicos, assim como, de comportamento geral, aptos a vários ambientes.

Segundo Vencovsky e Barriga (1992), a interação GxA é de natureza genética, mas não no sentido usual, e sim em decorrência de instabilidades das manifestações genóticas entre ambientes. Dessa forma, o comportamento dos genótipos em relação ao ambiente, tem merecido especial atenção por parte dos melhoristas, pois interfere nos processos de seleção. Por essa razão, torna-se importante o conhecimento preciso dessas estimativas, bem como sua utilização na determinação da estabilidade fenotípica dos diferentes genótipos.

A partir do fenômeno da interação GxA nos programas de melhoramento genético de plantas é possível executar uma seleção de genótipos com adaptação ampla ou específica (nesta tese, adaptação específica é tratada como adaptação preferencial), escolher locais de seleção, identificar o nível de estresse nos ambientes escolhidos, para as fases iniciais da seleção, e determinar o número ideal de ambientes e de genótipos a serem avaliados em cada fase da seleção (FOX et al., 1997).

A interação GxA constitui-se num dos maiores desafios dos programas de melhoramento de qualquer espécie, seja na fase de seleção ou na de recomendação de variedades. Entre as alternativas para minimizar esse problema está a escolha de genótipos com ampla adaptação e boa estabilidade (CRUZ; CARNEIRO, 2003).

O objetivo deste trabalho foi estimar parâmetros estatísticos e genéticos a partir de diferentes métodos, caracterização e seleção de genótipos de cana-de-açúcar quanto à produtividade, adaptabilidade e estabilidade em ensaios multi ambientais.

2. REVISÃO DE LITERATURA

2.1. A cana-de-açúcar (*Saccharum* spp.)

A cana-de-açúcar é uma espécie vegetal de grande importância para a agricultura brasileira e mundial. Segundo Daniels and Roach (1987), trata-se de uma espécie alógama, da família Poaceae) tribo Andropogoneae, gênero ***Saccharum***. Dentro do gênero *Saccharum*, ocorrem principalmente seis espécies: *S. officinarum* L. ($2n = 80$), *S. robustum* Brandes e Jeswiet ex Grassl ($2n = 60-205$), *S. barberi* Jeswiet ($2n = 81-124$), *S. sinense* Roxb. ($2n = 111-120$), *S. spontaneum* L. ($2n = 40-128$), e *S. edule* Hassk. ($2n =$

60-80). Como genomas de todas essas espécies podem estar participando de alguma forma das variedades modernas, considera-se que a cana-de-açúcar é uma das espécies cultivadas de maior complexidade genética.

O comportamento citogenético de espécies alopoliplóides é similar ao de indivíduos diplóides, ou seja, somente são formados bivalentes de cromossomos homólogos do mesmo genoma, visto que os genomas que deram origem aos alopoliplóides, em geral, diferem muito. Em cana-de-açúcar não se pode assumir herança diplóide. Entretanto, algumas evidências sugerem herança semelhante aos diplóides, e os desvios dessas suposições seriam não significativos. Dessa forma, as abordagens biométricas apresentadas nesta Tese são válidas para espécies diplóides e alopoliplóides e os modelos genéticos, portanto, podem ser empregados diretamente na análise de genótipos de cana-de-açúcar (HOGARTH, 1977; RESENDE, 2002, p.345).

Devido a alta complexidade da cana-de-açúcar, esta espécie é sujeita às interações que ocorrem entre o genótipo e o ambiente. A identificação e estudo da adaptabilidade e estabilidade dos genótipos na fase final de um programa de melhoramento de plantas, ou seja, ensaios em múltiplos ambientes, é o tema central desta tese.

2.2. Modelos Mistos: efeitos fixos e aleatórios, Máxima Verossimilhança Restrita (REML) e Melhor Predição Linear não-viesada (BLUP)

2.2.1. Efeitos fixos e aleatórios

Modelos estatísticos para dados são descrições matemáticas de como os dados podem ser concebivelmente produzidos. Modelos consistem ao menos de duas partes: i) Uma fórmula relativa às respostas para todas as variáveis explanatórias (por exemplo: efeitos), e ii) uma descrição da distribuição de probabilidade assumida para caracterizar a variação aleatória que afeta a resposta observada (LITTELL et al., 2006).

Há modelos de média que incluem somente uma distribuição assumida da variável resposta e uma função da média incluindo somente fatores de efeito que são tratados como constantes e esses efeitos constantes são chamados de efeitos fixos. Um efeito é chamado fixo se os níveis no estudo representam todos os possíveis níveis de um fator,

ou, pelo menos, todos os níveis sobre qual inferência está para ser feita. Isso também inclui modelos de regressão onde os valores observados da variável explanatória cobrem a região inteira de interesse sobre o fator que se quer inferir (LITTELL et al., 2006).

Fatores são aleatórios se eles são usados no estudo para representar uma amostra (idealmente, uma amostra aleatória) de um conjunto maior de níveis em potencial. O fator de efeito correspondente ao conjunto maior de níveis constitui uma população com uma distribuição de probabilidade, ou seja, um fator é considerado aleatório se seus níveis plausivelmente representarem uma população maior com uma distribuição de probabilidade. Efeitos aleatórios, então, representam uma amostragem verdadeira e assume-se haver uma distribuição de probabilidade (LITTELL et al., 2006). Pode-se dizer que os efeitos de genótipos são considerados como aleatórios porque os genes que levaram a formação de tais genótipos são uma amostra verdadeira de todos os genes que pertencem a população e, dessa forma, têm uma distribuição de probabilidade.

A decisão de que um fator é aleatório ou fixo nem sempre é fácil e pode haver controvérsias. O efeito de blocos e locais ilustram esse ponto. O uso de blocos em experimentos agrícolas frequentemente reflete variação em um campo, assim como num declive com um bloco numa faixa situada no topo dele, um bloco numa faixa logo abaixo e assim por diante até o fim deste declive. Pode-se argumentar que não há nada de aleatório sobre esses blocos. Porém, uma característica adicional de efeitos aleatórios é a permutabilidade (LITTELL et al., 2006).

2.2.2. Estimação de componentes de variância via Máxima Verossimilhança Restrita (REML)

O método REML foi desenvolvido por Patterson and Thompson (1971, APUD RESENDE, 2007; SORENSEN; GIANOLA, 2002), que apresentaram uma correção ao ML, eliminando o seu vício. No método REML, somente a porção da verossimilhança que é invariante aos efeitos fixos é maximizada. Assim, o REML mantém as demais propriedades do ML, é não viciado e permite também a imposição de restrições de não negatividade. Dessa forma, o REML é o procedimento ideal de estimação de componentes de variância com dados desbalanceados. Também Robertson (1962, APUD RESENDE, 2007), considerando a estimação de componentes de variância a partir de dados desbalanceados em modelos de classificação simples (única), derivou pesos ótimos para médias de família, os quais são, efetivamente, aqueles usados na estimação

REML. O procedimento REML é também denominado máxima verossimilhança marginal e, na Europa, por máxima verossimilhança residual.

Sorensen and Gianola (2002, p. 127-128) elucidaram o REML proposto por Patterson and Thompson (1971) e o colocaram da seguinte forma:

Com uma amostra de observações independentes n , segue-se a partir da equação

$$\sum_{i=1}^n (y_i - m)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - m)^2 \quad (1)$$

Onde na equação 1:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

A log-verossimilhança pode ser expressada como:

$$l(m|\sigma^2, y) = \text{constante} - \frac{1}{2\sigma^2} \left[\sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - m)^2 \right] \quad (2)$$

Maximizando (2) com relação ao m é equivalente a minimização da soma de desvios quadrados $\sum_{i=1}^n (y_i - m)^2$, então há uma relação com o critério dos quadrados mínimos neste caso. Levando a derivada de (2) com relação ao m , proporciona-se a forma linear de m :

$$\frac{dl(m|\sigma^2, y)}{dm} = \frac{n(\bar{y} - m)}{\sigma^2} \quad (3)$$

Igualando-se a equação 3 a zero e resolvendo para m , obtém-se, como máxima verossimilhança (ML) de m ,

$$\hat{m} = \frac{\sum_{i=1}^n y_i}{n} = \bar{y} \quad (4)$$

Consequentemente, o estimador de ML do valor esperado de uma distribuição normal com variância conhecida é a média aritmética das observações (o estimador teria sido o mesmo se a variância tivesse sido desconhecida, como demonstra-se mais tarde). Este é também o estimador de quadrados mínimos, mas há uma coincidência com a ML de m somente quando a normalidade é assumida, como neste exemplo. Diferenciando 3 em relação a m obtém-se $-n/\sigma^2$. Devido a esta derivada ser negativa, segue-se que $\hat{m}$ é um máximo.

Suponha-se, temporariamente, que m é uma variável aleatória real, e que σ^2 é um parâmetro desconhecido. Agora, integrando

$$L(m|\sigma^2, y) = \frac{1}{\sqrt{2\sigma^2}^n} \exp \left[-\frac{n(y_i - m)^2}{2\sigma^2} \right] \quad (5)$$

em relação a m obtém-se, para uma amostra de tamanho n ,

$$\int L(m|\sigma^2, y) dm = \frac{1}{(2\pi\sigma^2)^{(n-1)/2} \sqrt{n}} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2 \right] \quad (6)$$

que, dado os dados, é a função de σ^2 somente. (Os limites da integração são $\pm\infty$, mas isso é geralmente omitido a não ser que o contexto dite ao contrário). A expressão (6) é chamada de verossimilhança marginal ou integrada e tem a mesma forma da chamada verossimilhança restrita, proposta por Patterson and Thompson (1971). De fato, quando σ^2 é um parâmetro desconhecido, a maximização de (6) em relação a σ^2 produz

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1},$$

que é o estimador da Máxima Verossimilhança Restrita (REML) de σ^2 (SORENSEN; GIANOLA, 2002).

Os métodos ML e REML requerem solução iterativa, pela não linearidade das equações, fato que dificulta a derivação de estimadores explícitos. Assim, os componentes de variância iteram nas equações de modelo misto do BLUP até convergência para um valor adequado. Na maioria das vezes, a estimação dos componentes de variância e a predição de valores genéticos são realizados, simultaneamente, pelo procedimento REML/BLUP. Neste caso, o procedimento BLUP é denominado BLUP empírico (RESENDE, 2007).

Os métodos de máxima verossimilhança permitem, sob certas condições, levar em conta os efeitos da seleção. É necessário que todas as informações que tenham contribuído para a seleção sejam incluídas na análise (bem como o parentesco entre os indivíduos), exceto se essas informações não forem correlacionadas com o caráter sob análise (MEYER, 1989). De maneira geral, argumentos e evidências indicam que inferências pontuais realizadas a partir de funções de verossimilhança não são afetadas por algumas formas de seleção (GIANOLA et al., 1989). Essa propriedade fez com que o método REML se tornasse padrão para a estimação de componentes de variância em programas de melhoramento genético.

Kennedy and Sorensen (1988) mostram também que o método REML elimina o

vício devido a mudanças nas frequências alélicas pela seleção, pelo uso da matriz de parentesco completa (A) quando se trabalha com ela. Assim, torna-se possível a obtenção de componentes de variância para uma população base não selecionada e a predição de valores genéticos de indivíduos de quaisquer gerações é realizada com precisão. O uso da matriz A completa leva em consideração as alterações na variância genética devidas à endogamia e ao desequilíbrio de ligação, as quais resultam da seleção e, considera também, a tendência genética ou ganho genético realizado.

Os métodos REML e ML exigem normalidade para que os estimadores tenham propriedades desejáveis. Entretanto, tais estimadores podem ser robustos aos desvios da normalidade, gerando estimativas razoáveis mesmo quando a forma da distribuição não é especificada (HARVILLE, 1977). Para dados balanceados, o estimador REML (negligenciando a suposição de normalidade e a restrição de estimativas REML ao espaço paramétrico) é idêntico ao estimador LS (análise de variância), o qual apresenta as propriedades desejáveis BUE (melhor estimador não viciado) e BQUE (melhor estimador quadrático não viciado) sob normalidade (SEARLE et al., 1992).

Diferentemente de estimadores de ML, os estimadores REML maximizam somente a porção da verossimilhança que não depende dos efeitos fixos. Neste sentido, o REML é uma versão restrita da ML. A eliminação do viés pelo REML é análoga à remoção do viés que surge na estimativa do componente de variância quando o desvio quadrado médio é dividido pelo grau de liberdade ao invés de pelo tamanho amostral. O REML nem sempre elimina todos os vieses na estimação de parâmetros, uma vez que muitos métodos para obtenção de estimativas de REML não podem retornar estimativas negativas de um componente de variância. Porém, esta fonte de viés também existe com ML, então o REML é claramente o método preferido para analisar grandes conjuntos de dados com estruturas complexas. No caso ideal de um delineamento inteiramente balanceado, o REML produz componentes de estimativas de variância que são idênticos àqueles obtidos pelo análise clássica de variância (LYNCH; WALSH, 1998).

2.2.3. Melhor Predição Linear não-viesada (BLUP)

Um modelo misto contém efeitos tanto fixos quanto efeitos aleatórios. Uma das principais características que distingue esta metodologia de métodos de modelos lineares convencionais é a habilidade para estimar "valores reais de variáveis aleatórias", ou seja,

efeitos aleatórios específicos ou funções lineares de efeitos aleatórios (LITTELL et al., 2006). Por exemplo, Henderson (1963) desenvolveu um procedimento para predição de valores genéticos de touros em experimentos genéticos com animais. Para fazê-lo, ele usou estimativas de efeitos aleatórios de touros em um modelo misto. Henderson chamou este procedimento de "Best Linear Unbiased Prediction", ou BLUP. Harville (1976) mostrou que o procedimento de Henderson tinha uma base teórica válida: Ela podia ser justificada como uma extensão do Teorema de Gauss-Markov. Mais recentemente, o uso de preditores baseado em efeitos aleatórios estimados ganhou ampla aceitação e tem sido aplicado em muitos problemas estatísticos.

Tem-se alguns exemplos sobre a adoção do BLUP como estimativa de valores genéticos em culturas perenes, semi-perenes e anuais, tais como: Bernardo (1996); Souza et al. (2000); Paiva et al. (2002); Eagles and Moody (2004); Missio et al. (2004, 2005); Reis et al. (2005); Sebolai et al. (2005); Bauer et al. (2006); Crossa et al. (2006); Flachenecker et al. (2006a e 2006b); Bastos et al. (2007); Piepho et al. (2008); Nunes et al. (2008); Farias Neto et al. (2009); Pedrozo et al. (2009); Verardi et al. (2009); Arantes et al. (2010); Borges et al. (2010a); Viana et al. (2010a e 2010b); Mi et al. (2011), Arantes et al. (2013) e etc.

Segundo Piepho et al. (2008), tanto o desenvolvimento de novas variedades quanto as suas recomendações de lançamento exigem uma seleção para ser feita entre um conjunto maior de genótipos candidatos, então a estimação de valores genéticos está no centro de qualquer esforço para o melhoramento. Valor genético refere-se a decomposição do valor genotípico para efeitos de componentes com uma interpretação genética, decompondo-se em efeitos aditivos e não aditivos (Dominância, Epistasia). O ponto final do objetivo de um programa de melhoramento genético geralmente é a predição dos valores genotípicos de um genótipo candidato.

Bernardo (2010) afirma que dados de desempenho de estágios iniciais de seleção em uma população de melhoramento são geralmente balanceados. Isso significa que famílias, linhagens, test-crosses, ou híbridos são avaliados no mesmo grupo de locais e anos. Ter dados balanceados simplifica altamente a tarefa de selecionar os melhores genótipos numa população de melhoramento. Suponha que um melhorista de soja desenvolva 20 linhagens (inbred) recombinantes da população 1 e 15 linhagens recombinantes da população 2. As 20 linhagens recombinantes da população 1 são avaliadas em quatro locais, enquanto que as 15 linhagens recombinantes da população 2 são avaliadas em seis diferentes locais. As melhores linhagens de cada população

podem ser selecionadas com base em seus desempenhos médios nos respectivos testes de performance.

Mas suponha que o melhorista agora deseja comparar as 20 linhagens recombinantes da população 1 com as 15 da população 2. Um problema com esta comparações é que os dados agora estão desbalanceados: nem todos as 35 linhagens foram avaliadas em todos os 10 locais. Diferenças entre grupo de ambientes são confundidos com diferenças entre grupos de linhagens recombinantes. Uma comparação direta, portanto, não pode ser feita baseando-se em médias simples das linhagens (BERNARDO, 2010).

De acordo com Bernardo (2010), BLUP é um procedimento geral que permite comparações entre genótipos desenvolvidos de diferentes populações de melhoramento e avaliados em diferentes conjuntos de ambientes. O procedimento BLUP é útil por duas razões:

1) Permite a análise de dados desbalanceados acumulados de testes de performance. Mais dados se tornam disponíveis para os melhores genótipos porque são testados mais amplamente enquanto os genótipos menos superiores são descartados. O procedimento de BLUP lida com dados desbalanceados enquanto considera as diferenças na quantidade de dados disponíveis para cada genótipo.

2) O procedimento de BLUP explora informações de parentes, por meio disso, maximiza o uso dos dados disponíveis quando compara os genótipos. Suponha que um melhorista deseja comparar duas linhagens, A e B. Esta comparação pode ser feita baseando-se no desempenho de A e B sozinhos. Porém, esta comparação se torna mais precisa quando a performance dos parentes de A e dos parentes de B também são consideradas. O conceito da informação dos parentes é talvez melhor percebido com um exemplo de gado leiteiro. Touros carregam genes que afetam a produção de leite mas por razões óbvias um touro não pode ser diretamente avaliado por sua produção de leite. O valor genético de um touro para produção de leite, porém, pode ser predito da produção de leite do seu parente feminino (FALCONER; MACKAY, 1996). Informações do parental da cultura a partir do pedigree é altamente importante e facilmente obtida em programas de melhoramento.

C.R. Henderson começou seu trabalho pioneiro no BLUP em meados de 1940 (HENDERSON et al., 1959; HENDERSON, 1963, 1975). O procedimento BLUP primeiro usado em 1970 para avaliar cerca de 1200 touros num programa de inseminação artificial (FREEMAN; HENDERSON, 1991; SCHAEFFER; HENDERSON, 1991). Desde então o

BLUP tem sido usado rotineiramente no melhoramento animal. O procedimento BLUP não foi usado em plantas até meados dos anos de 1990 (BERNARDO, 1994; PANTER; ALLEN, 1995). O nosso uso do termo BLUP na verdade refere-se à junção do uso de ambos BLUP e "Best Linear Unbiased Estimation", ou BLUE. Estes dois procedimentos diferem em que efeitos são considerados fixos no BLUE e aleatórios no BLUP:

1) Efeitos fixos são estimados no BLUE. Estes efeitos fixos são constantes ao invés de variáveis aleatórias. Exemplos de efeitos fixos são média geral, efeitos de diferentes tipos de solos ou de um transgene. Em melhoramento de plantas um ambiente individual é considerado aleatório mas a média de um grupo de ambientes usada, no caso de testes de performance, é considerada um efeito fixo (BERNARDO, 1996). Diferenças entre as médias dos conjuntos de ambientes são tratadas como fatores de impedimento que precisam ser levados em consideração antes de comparar genótipos. No nosso exemplo com soja, suponhamos quatro locais nos quais as linhagens da população 1 foram avaliadas havendo menos precipitação do que nos seis locais nos quais as linhagens da população 2 foram avaliadas. As observações precisam ser corrigidas para quaisquer efeitos de precipitação antes de comparar a performance de uma linhagem da população 1 e uma linhagem da população 2.

2) Efeitos Aleatórios são preditos por BLUP. Efeitos aleatórios tem uma estrutura de covariância enquanto que efeitos fixos não a tem. Por exemplo, um parental passa metade do valor genético para suas progênies, que causa uma covariância ou correlação entre os dois. Efeitos genéticos tais como: valores genéticos e desvio de dominância, efeitos de testcross, e o efeito de habilidade de combinação geral e específica no melhoramento de populações tem uma estrutura de covariância e são, portanto, considerados aleatórios.

A presença de ambos os efeitos fixos e aleatórios leva-nos aos modelos mistos. BLUP e BLUE, considerados procedimentos, são, na verdade, propriedades estatísticas de predição e estimação. "Best" indica que a variância da amostra do que está sendo estimado ou predito é minimizada. "Linear" significa que as estimativas ou predições são funções lineares das observações. "Unbiased" no BLUE indica que os valores esperados das estimativas são iguais aos seus valores verdadeiros. Efeitos genéticos são definidos como tendo média zero, e "Unbiased" no BLUP indica que as predições tem uma expectativa de zero.

O conhecimento dos valores verdadeiros das variâncias e covariâncias dos efeitos aleatórios é exigido no BLUP. Mas os valores verdadeiros das variâncias genéticas e não

genéticas são desconhecidas, então qualquer implementação do BLUP usando estimativas destas variâncias é uma aproximação. Na prática, BLUP em programas de melhoramento genético frequentemente envolve a predição simultânea de efeitos genéticos e a estimação de variâncias genéticas e não genéticas. Com exceção de delineamento experimental de um fator, delineamentos de cruzamento são raramente usados por melhoristas porque exigem experimentos para tal. A metodologia de BLUP permite a estimação de variâncias genéticas dos mesmos dados que são rotineiramente gerados nos programas de melhoramento, por isso evita-se a necessidade por delineamentos de cruzamento (BERNARDO, 2010).

Uma propriedade principal do BLUP é o encolhimento em direção a média. Este encolhimento é mais encontrado como "shrinkage", do termo em inglês, mesmo em trabalhos em português e será tratado dessa forma nesta tese. Littell et al. (2006) definiram que o shrinkage para a variável resposta para um dado nível de um fator aleatório é a combinação ponderada da média geral, baseado nos efeitos fixos e na média ordinária para o nível do fator aleatório. O shrinkage depende da variância entre os níveis do fator aleatório e da variância dos valores da variável observada para cada nível do fator aleatório. No caso dos modelos de componentes de variância, pode-se escrever:

$$b_i = \alpha_i \mu + (1 - \alpha_i) \mu_i \quad (1)$$

Tem-se a seguinte notação da equação acima:

b_i : shrinkage médio para o nível i do fator aleatório;

α_i : fator ponderado do nível i calculado por $\sigma_i^2 / (\sigma_{\text{random factor}}^2 + \sigma_i^2)$;

σ_i^2 : variância para o nível i do fator aleatório, calculado por $\sigma_{\text{error}}^2 / n_i$;

$\sigma_{\text{random factor}}^2$: variância dos efeitos aleatórios associada ao fator aleatório;

μ : média geral dos valores das respostas;

μ_i : média dos valores das respostas para o nível i do fator aleatório.

Quando calcula-se o shrinkage médio, percebe-se o uso das estimativas da amostra para as quantidades na equação 1. Esta equação mostra como a quantidade do shrinkage nos BLUPs é uma função da variância dos valores da resposta observada dentro de um nível do fator aleatório e a variância dos efeitos aleatórios. Em casos extremos em que não haja variância entre os níveis do efeito aleatório ($\sigma_{\text{random factor}}^2 = 0$), há um shrinkage total, e a estimativa da média do shrinkage para cada nível do fator aleatório é igual a média geral. Ou seja, não há diferença na predição média para cada

nível do fator aleatório.

Como a equação 1 mostra, a quantidade de "shrinkage" será uma função da variância do efeito aleatório e da variância total dos dados observados para um dado nível do fator aleatório. Se houver muita variabilidade entre as médias da variável resposta para os diferentes níveis do fator aleatório, haverá menos shrinkage do que se houvesse menos variabilidade. No caso extremos quando $\sigma^2_{\text{random factor}} = 0$, o efeito estimado de cada nível do fator aleatório será também zero. No caso onde $\sigma^2_{\text{random factor}} > 0$, a extensão para qual o BLUP para um particular efeito aleatório é "shrunk" (encolhido) depende da quantidade de informação disponível para predição daquele efeito aleatório. Por exemplo, se predizer BLUPs para grupos amostrados aleatoriamente, a predição do efeito do grupo aleatório será encolhido menos por grupos com tamanhos amostrais grandes do que para grupos com tamanhos amostrais menores, porque não tem-se mais informação sobre grupos maiores. Finalmente, segundo Searle et al. (1992), sob certas hipóteses bastantes gerais, BLUP maximiza a correlação dos valores genotípicos reais e os valores genotípicos preditos.

Segundo Piepho et al. (2008), BLUP parece não ter ganhado a mesma popularidade como no melhoramento genético animal, onde este é um procedimento padrão. Uma razão pode ser que nos programas de melhoramento genético de plantas geralmente produzem relativamente uma grande quantidade de dados fenotípicos por genótipo a partir de uma sucessão de experimentos, então BLUE e BLUP frequentemente fornecem resultados semelhantes. Isto pode levar a uma falsa impressão que a escolha do estimador não é importante.

2.3. Inferência Bayesiana de componentes de variância

A inferência Bayesiana, ao contrário da inferência clássica, ou seja, a frequentista, leva em conta o conceito de probabilidade. Na prática, a maior diferença é tentar medir o grau de incerteza que se tem sobre a ocorrência de um determinado evento do espaço amostral, utilizando distribuições de probabilidades a priori e a amostral (verossimilhança). Na inferência Bayesiana, a incerteza sobre os parâmetros desconhecidos associa-se uma distribuição de probabilidade (GIANOLA; FERNANDO, 1986), enquanto que, na inferência frequentista, os parâmetros são valores fixos ou constantes, aos quais não se associam a qualquer distribuição (BLASCO, 2001). Sob o enfoque Bayesiano, os efeitos de

tratamentos são sempre considerados aleatórios, de forma que estimadores que promovem "shrinkage" são sempre utilizados, mesmo com números de tratamentos menores que 10 (RESENDE, 2004).

Segundo Resende (2000 e 2007), a inferência estatística Bayesiana baseia-se na distribuição condicional do parâmetro (θ) dado o vetor de dados (y), ou seja, na distribuição a posteriori do parâmetro dadas as observações fenotípicas. O Teorema de Bayes, definido em termos de densidades de probabilidade, tem a seguinte formulação para a distribuição de uma variável aleatória contínua:

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{\int_R f(y|\theta)f(\theta) d\theta}$$

θ : Vetor de parâmetros

y : Vetor de dados ou de informações obtidas por amostragem

$f(\theta|y)$: distribuição condicional de θ dado y , ou distribuição a posteriori (que é a base da estimação e predição bayesiana).

$f(y|\theta)$: função densidade de probabilidade da distribuição condicional de uma observação (y) dado θ (denominada função de verossimilhança ou modelo para os dados).

$f(\theta)$: função densidade de probabilidade da distribuição a priori, que é também a densidade marginal de θ . Esta função denota o grau de conhecimento acumulado sobre θ , antes da observação de y .

$f(y|\theta)f(\theta)$: função densidade conjunta de y e θ .

$$f(y) = \int_R f(y, \theta) d\theta = \int_R f(y|\theta)f(\theta) d\theta = E_\theta[f(y|\theta)]$$

$f(y)$: distribuição marginal ou preditiva de y com respeito a θ , onde R é a amplitude da distribuição de θ . E_θ significa esperança com respeito à distribuição de θ . (A integração da distribuição conjunta, no espaço paramétrico, produz a marginal de y).

Verifica-se, então, que a distribuição a posteriori é proporcional à verossimilhança x priori, ou seja, a função de verossimilhança conecta a priori a posteriori usando para isto os dados do experimento (observações). Dessa forma, a distribuição a posteriori contempla o grau de conhecimento prévio sobre o parâmetro (θ) e também as informações adicionais propiciadas pelo experimento (y) (RESENDE, 2000 e 2007).

Segundo RESENDE (1997), componentes de variância são estimados pelo método

REML e utilizados na predição de valores genéticos e cálculo do ganho genético. Entretanto, através deste procedimento, a distribuição e variância dos estimadores não são conhecidas e, assim, questões referentes à efetividade da seleção não podem ser respondidas com rigor. O método bayesiano de estimação dos componentes de variância, implementado via simulação estocástica, como por exemplo através do procedimento denominado amostrador de Gibbs, propicia uma descrição mais completa sobre a confiabilidade dos parâmetros genéticos do que o método REML. A abordagem bayesiana (conhecimento da distribuição a posteriori dos parâmetros) possibilita a construção de intervalos de confiança (ou de probabilidade) exatos para as estimativas dos parâmetros genéticos, estando a confiabilidade (precisão) dos parâmetros associadas com a quantidade e precisão das informações utilizadas. Dessa forma o método amostrador de Gibbs propicia estimativas mais precisas de componentes de variância, parâmetros genéticos, valores genéticos, ganhos genéticos e menores desvios padrões que o método REML.

2.3.1. Amostrador de Gibbs

Em genética quantitativa, para implementação prática da análise bayesiana, uma das maiores dificuldades técnicas é a marginalização. A obtenção de distribuições marginais por processos analíticos é praticamente impossível (SORENSEN, 1996). Assim, a obtenção da distribuição marginal a posteriori (marginalização da distribuição conjunta a posteriori) tem sido obtida pelo método da amostragem de Gibbs através da amostragem e atualização das distribuições condicionais (RESENDE, 2004).

Casella and George (1992) ilustram o amostrador de Gibbs da seguinte forma: Suponha que tem-se uma densidade conjunta dada como $f(y, \theta_1, \dots, \theta_p)$, e há interesse na obtenção de características da densidade marginal

$$f(y) = \int \dots \int f(y, \theta_1, \dots, \theta_p) d\theta_1 \dots d\theta_p \quad (\text{Equação 2})$$

assim como a média ou variância. Talvez a mais natural e simples abordagem seria calcular $f(y)$ e usá-la para obter a característica desejada. Porém, há muitos casos onde as integrações na equação 2 são extremamente difíceis de se obter, tanto quanto analiticamente quanto numericamente. Em tais casos o amostrador de Gibbs propicia um método alternativo para obter a $f(y)$, ou seja, a função de densidade marginal.

Ao invés de computar ou aproximar a $f(y)$ diretamente, o amostrador de Gibbs permite efetivamente gerar uma amostra $Y_1, \dots, Y_m \sim f(y)$ sem requerer a função de densidade marginal $f(y)$. Pela simulação de uma grande amostragem suficiente, a média, variância, ou qualquer outra característica de $f(y)$ pode ser calculada para o grau desejado de acurácia.

É importante perceber que, no efeito, o resultado final de quaisquer cálculos, embora baseados em simulação, são quantidades da população. Por exemplo, para calcular a média de $f(y)$, pode-se usar $(1/m) \sum_{i=1}^m Y_i$, e o fato que

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m Y_i = \int_{-\infty}^{\infty} y f(y) dy = EY$$

Portanto, pegando uma quantidade suficiente de m , qualquer característica da população, até mesmo a densidade, pode ser obtida para qualquer grau de acurácia (CASELLA; GEORGE, 1992).

Após a obtenção de uma grande quantidade m de amostras, conseqüentemente após um grande número de iterações, a sequência de valores gerados pelo amostrador de Gibbs converge para uma distribuição estacionária igual a $f(\theta|y)$, sendo que cada valor de θ , obtido pelo amostrador de Gibbs após a convergência, é um valor simulado da distribuição conjunta de seus elementos (GEMAN; GEMAN, 1984). Desse modo, a sequência de valores gerados das amostras aleatórias, extraídas das condicionais, se aproxima da distribuição de equilíbrio, ou seja, são equivalentes às amostras aleatórias da distribuição conjunta e/ou marginal a posteriori (CASELLA; GEORGE, 1992).

O amostrador de Gibbs é um processo iterativo de amostragem de uma cadeia de Markov. Tais cadeias estão inseridas no contexto da teoria dos processos estocásticos definido como a parte dinâmica da teoria de probabilidades, onde se estuda uma coleção de variáveis aleatórias, com respeito a sua interdependência e comportamento limite. Para a inferência bayesiana, é de maior relevância o estudo do comportamento assintótico da cadeia, quando o número de iterações tende a ∞ , uma vez que a inferência deve ser baseada na distribuição (a posteriori) estacionária, ou seja, em equilíbrio (GAMERMAN, 1997 apud RESENDE, 2004).

Há diferentes formas para se analisar o comportamento assintótico da cadeia de Markov, ou seja, detectar a convergência da cadeia de Gibbs, assim, vários testes, os quais verificam a estacionalidade da cadeia, o descarte inicial "burn-in", o tamanho efetivo da amostra, entre outros, são descritos em diversas literaturas (GELMAN; RUBIN, 1992;

GEWEKE, 1992; RAFTERY; LEWIS 1992a,b).

2.4. Interação genótipo x ambiente, adaptabilidade e estabilidade

Caracteres quantitativos, especialmente aqueles que são controlados por muitos genes, são em geral altamente influenciados por fatores ambientais e demonstram uma variação contínua. A maior parte das características agronomicamente e economicamente importante, tal como a produção de grãos e produção de carne, são quantitativos ou multigênicos na natureza. Devida a presença universal da interação genótipo x ambiente (GxA) de características quantitativas, avaliações genotípicas em, pelo menos, estágios avançados de melhoramento de plantas são executados em ambientes múltiplos em experimentos planejados. Os principais objetivos destes experimentos são (1) comparar a performance genotípica em duas inferências básicas - por exemplo: ampla inferência, a performance geral de um genótipo (através dos ambientes), e ambiente específico ou inferência restrita, a performance de um genótipo dentro de um ambiente específico; (2) estimar o componente de interação GxA para calibrar a herdabilidade e seu impacto na seleção, selecionar sites testes e macro-ambientes, identificar genótipos especificamente adaptados para ambientes alvos, e estabelecer objetivos de melhoramento genético (YAN; KANG, 2003).

Há várias técnicas para se avaliar a interação genótipo x ambiente. Entre as mais utilizadas, destacam-se as que procuram relacionar a produção (Y) de um genótipo em função de um índice ambientais (X). Teoricamente, demonstra-se que se X e Y são binormais, o valor esperado de Y em função dos observados valores de X é uma reta, cujo coeficiente angular depende da correlação entre X e Y e dos respectivos desvios padrões (HOGG; CRAIG, 1965).

Diferentes metodologias para avaliar a adaptabilidade e a estabilidade tem sido desenvolvidas e, ou, aprimoradas. Tais procedimentos se baseiam em análise de variância, regressão linear, regressão não linear, análises multivariadas e estatísticas não paramétricas (BASTOS et al., 2007).

A estabilidade refere-se à capacidade de os genótipos mostrarem um comportamento constante (estabilidade fenotípica).

Adaptabilidade é a capacidade de responder positivamente aos estímulos dos ambientes. Em Bernardo (2010) este conceito é também chamado de estabilidade

dinâmica.

Quando um genótipo não possui um comportamento previsível em função dos ambientes ele pode eventualmente ter resposta favorável a ambientes específicos (adaptabilidade preferencial ou específica para determinados ambientes); o que sugere, para esse caso, seleção regional ou para locais específicos.

Por essas definições, necessariamente os desvios em relação ao modelo proposto (linear ou não) devem ser não significativos. No caso contrário (desvios do modelo significativos), há falta de adaptabilidade geral aos ambientes.

O método de Lin and Binns (1988) é bastante usado para estudar a interação GxA e baseia-se em análise não paramétrica. Nesta metodologia o desempenho dos acessos é quantificada pelo índice de estabilidade P_i , que corresponde ao quadrado médio da distância entre a média de um acesso para um dado ambiente e a resposta máxima para o mesmo ambiente, em todos os ambientes avaliados. Dessa forma, o quadrado médio menor indica uma superioridade geral do genótipo em questão, pois quanto menor o valor de P_i , menor será o desvio em torno da produtividade máxima; assim, maior estabilidade está relacionada, obrigatoriamente, com alta produtividade (DAHER et al., 2003).

O método de Eberhart and Russell (1966), baseado em regressão linear, é um dos métodos mais utilizados para estudar a adaptabilidade e estabilidade dos genótipos. Neste processo um índice ambiental é calculado através da subtração entre a média do ambiente (média de todos os genótipos neste ambiente) menos a média geral (média geral de todos os genótipos em todos os ambientes). Usando o índice ambiental como abscissa X e o valor observado do genótipo em determinado ambiente como ordenada Y, utiliza-se uma regressão linear para cada genótipo testado ($Y = \alpha + bX$). O valor de b (ângulo) estima a adaptabilidade do genótipo. Quando $b < 1$, pode-se afirmar que o genótipo responde pouco à melhoria ambiental, $b = 1$ corresponde a genótipos que respondem de forma mediana à melhoria ambiental e $b > 1$ corresponde a genótipos que respondem de forma acentuada à melhoria de ambiente e são mais indicados para ambientes superiores. Neste método, a estabilidade é estimada pelo somatório dos desvios ($MS-D_i$), da resposta em relação ao predito pela reta.

No contexto dos modelos mistos, um método para ordenamento de genótipos simultaneamente por seus valores genéticos (produtividade) e estabilidade, refere-se ao procedimento BLUP sob médias harmônicas (RESENDE, 2002). Neste caso, o vetor de dados (y) deve ser trabalhado como a recíproca dos dados observados, ou seja, $(1/y)$. Isto produz resultados que são funções $(1/H)$ da média harmônica (H) dos dados. Quanto

menor for o desvio padrão do comportamento genotípico através dos locais, maior será a média harmônica de seus valores genotípicos através dos locais. Assim, a seleção pelos maiores valores da média harmônica dos valores genotípicos (MHVG) implica simultaneamente seleção para produtividade e estabilidade. Adicionalmente, o uso da transformação $1/y$ considera também a instabilidade dentro dos locais.

A seleção considerando simultaneamente a produtividade, estabilidade e adaptabilidade, no contexto dos modelos mistos, pode ser feita por meio da estatística da Média Harmônica da Performance Relativa dos Valores Genotípicos (MHPRVG) preditos, apresentando as seguintes vantagens, quando comparado com outros métodos, segundo Resende (2004): (i) considera os efeitos genotípicos como aleatórios; (ii) permite lidar com heterogeneidade de variâncias; (iii) permite lidar com dados não balanceados; (iv) permite lidar com delineamentos não ortogonais; (v) considera os erros correlacionados dentro de locais; (vi) fornece valores genéticos já descontados da instabilidade; e (vii) gera resultados na própria grandeza ou escala do caráter avaliado.

2.4.1. Regressão Linear Fenotípica Conjunta (EBERHART; RUSSELL, 1966)

A Regressão Linear Fenotípica Conjunta proposta por Eberhart and Russell (1966) é uma metodologia de análise de estabilidade bastante utilizada em programas de melhoramento de plantas. Este tipo de análise visa examinar a reação de um genótipo, relativa aos outros genótipos, para diferentes ambientes. A análise de estabilidade permite a identificação de genótipos que são estáveis ou instáveis (BERNARDO, 2010).

O conceito de estabilidade implica que alguma medida que distingue um ambiente do outro seja necessária, um índice que possa representar o ambiente. Idealmente este Índice Ambiental (I_j) deveria ser baseado em fatores ambientais que afetam a performance dos genótipos, tais como propriedades do solo, fatores climáticos, ou estresses bióticos ou abióticos (EBERHART; RUSSELL, 1966; LIN et al., 1986 APUD BERNARDO, 2010). Mas até que tais índices são desenvolvidos, o efeito de j -ésimo ambiente pode servir como um I_j útil na análise de estabilidade de genótipos. Entretanto os genótipos devem ser cultivados numa grande variação de ambientes e estes ambientes devem ser distribuídos uniformemente nesta variação. De acordo com Bernardo (2010) a análise de estabilidade é menos significativa se, por exemplo, altos e baixos valores do efeito ambiental estão presentes mas valores intermediários destes

efeitos estão ausentes.

O uso da regressão linear para identificar genótipos estáveis foi primeiramente proposto por Yates and Cochran (1938) que tiveram a ideia de divisão das interações genótipo específicas em um componente explicado pela performance da média da população e um componente residual. Esta abordagem, que tornou-se a análise de regressão conjunta, possui algumas dificuldades estatísticas e foi bastante omitida até o uso por Finlay and Wilkinson em 1963, aplicando-a a variedades de cevada. Três anos depois Eberhart and Russell (1966) propuseram o mesmo método mas a regressão foi baseada em um Índice Ambiental calculado ao invés de se usar a performance média dos genótipos no ambiente. Desde então tem sido utilizada amplamente na análise de cultivo de variedades em vários experimentos (BERNARDO, 2010; LYNCH; WALSH, 1998).

Há muitos métodos que são propostos para análise estatística da interação genótipo x ambiente (GxA), e muitos trabalhos tem sido feitos tanto por estatísticos interessados na não aditividade, em geral, quanto para agrônomos, melhoristas e geneticistas, que estão interessados nas interações genótipo x ambiente. Uma grande quantidade de trabalhos foram publicados e nesta tese há uma tentativa de incluí-los numa breve revisão.

Fisher and Mackenzie (1923) publicaram um artigo de importância histórica que deu uma ideia indiretamente da interação GxA e citaram que as produções de diferentes variedades sob tratamento de adubações diferentes são melhores ajustadas pelo produto da fórmula do que pela soma da fórmula. Porém este trabalho permaneceu ignorado por muito tempo. A técnica de componente de variância foi seguida por muitos, por exemplo, Miller, Williams e Robinson (1959) que consideraram a ideia para discutir a interação GxA na cultura do algodão. Alguns trabalhos aplicaram a transformação de dados para uma escala de produção sem interação. Bartlett (1947) e Tukey (1949) sugeriram tal procedimento estatístico que foi aplicado para superar os problemas levantados pela presença da interação GxA. Uma outra abordagem, usada por geneticistas, foi entender a causa da interação GxA em termos de parâmetros genéticos biométricos. A abordagem foi então inteiramente diferente daquela dos melhoristas e agrônomos, os geneticistas gostariam de minimizar os efeitos da interação GxA nos seus campos. Em estudos genéticos, Morely and Mather (1958) discutiram como a interação GxA influencia as variâncias e co-variâncias.

Baseados nos estudos de Yates e Cochran (1938), Finlay e Wilkinson (1963) redescobriram um método similar e propuseram que a performance dos genótipos em

ambientes diferentes pode ser calculada para uma grande amplitude pela suas regressões sobre o índice do ambiente. Foi ainda estendido por Eberhart and Russell (1966) e Perkins and Jinks (1968a,b). Principalmente cinco abordagens de regressão tem sido utilizadas para descrever e analisar a interação GxA. A primeira abordagem, adotada por Finlay e Wilkinson (1963), Row and Andrew (1964), Eberhart and Russell (1966), Bucio Alanis (1966), Bucio Alanis and Hill (1966) e Breese (1969), é puramente estatística, que envolve o uso inválido estatisticamente das estimativas de não independência das variáveis. A segunda abordagem, adotada por Perkins and Jinks (1968a,b) e Baker (1969), e a terceira abordagem, derivada pelo Freeman and Perkins (1971), relatam os componentes na análise de regressão para o modelo genético biométrico básico. A quarta abordagem por Shukla (1972), e a quinta abordagem por Hardwick and Wood (1972), Perkins (1972) e Shukla (1983), são baseadas no conceito de regressão múltipla sobre variáveis externas como meios de encontrar as causas fisiológicas envolvidas da regressão observada sobre a média ambiental. Entretanto, em todas essas abordagens de procedimentos de estimação foram derivadas do uso do princípio dos Quadrados Mínimos.

A primeira abordagem no qual o valor fenotípico é regressado sobre o Índice Ambiental proporciona duas medidas de sensibilidade para mudança no ambiente em termos do coeficiente de regressão e o desvio quadrado da média da regressão linear. Medidas similares são derivadas na segunda abordagem pela regressão do componente GxA sobre o componente ambiental aditivo.

O modelo de Eberhart and Russell é idêntico ao de Finlay e Wilkinson, no qual os coeficientes de regressão médio entre todos os genótipos são igual a um, ou seja, a esperança do coeficiente de regressão é igual a um.

Segundo Lynch and Walsh (1998) a aplicação da equação do modelo de Eberhart and Russell (1966), $Y_{ij} = \beta_{oi} + \beta_{1i}I_j + \delta_{ij}$, para todos os genótipos em uma análise de interação genótipo x ambiente proporciona uma base para ordenar os genótipos com respeito às suas responsividades à mudança ambiental. $\beta_{1i} = 0$ implica que a resposta média do genótipo i para o ambiente é a mesma que a resposta média da população de genótipos na análise. $\beta_{1i} > 0$ implica em uma resposta média mais forte, enquanto que $-1 < \beta_{1i} < 0$ implica que o genótipo tende a responder à mudança ambiental numa direção contrária à média dos genótipos na população.

É importante notar que β_{1i} é somente uma medida de uma tendência do genótipo

para responder à mudança ambiental de maneira paralela à média da população. Dois genótipos com idênticos coeficientes β_{1i} podem, de fato, ter um conjunto de ambientes (I_j) bastante diferentes (LYNCH; WALSH, 1998).

2.4.2. Média Harmônica da Performance Relativa do Valor Genotípico (RESENDE, 2004)

Os procedimentos de análise de interação genótipo x ambiente evoluíram da tradicional ANOVA conjunta de experimentos, passando pelos métodos de estudo da estabilidade e adaptabilidade fenotípica baseados em análise de regressão, pelos métodos não paramétricos para estabilidade e adaptabilidade e pelos modelos multiplicativos (AMMI) para os efeitos de interação. Tais procedimentos apresentam limitações para lidar com dados desbalanceados, delineamentos experimentais não ortogonais (blocos incompletos) e com a heterogeneidade de variâncias entre os vários locais de experimentação, situações estas corriqueiras na experimentação de campo. Além do mais, tais metodologias assumem, em geral, que os efeitos de tratamentos genéticos são fixos, o que é desvantajoso e incoerente com a prática simultânea da estimação de componentes de variância e parâmetros genéticos (tais quais a herdabilidade) realizada com base nestes experimentos (RESENDE, 2007).

Outro aspecto, ainda segundo Resende (2007), refere-se à escolha do procedimento a ser aplicado, dentre os vários disponíveis. Alguns procedimentos conduzem a resultados similares, outros possuem propriedades estatísticas superiores e alguns permitem interpretações mais simples dos resultados. Assim, a escolha a priori do método a aplicar é tarefa difícil devido a vários métodos não serem prontamente comparáveis, pelo menos formalmente, pois envolvem conceitos diferentes. Neste sentido, Piepho (1999 APUD RESENDE, 2007) propôs o uso da metodologia de modelos mistos via REML para a comparação entre os vários procedimentos, tais quais o de Finlay e Wilkinson, o de Eberhart and Russell, o de Shukla, o de Lin et al. e uma versão AMMI considerando os efeitos de locais como aleatórios. Segundo Piepho (1999 APUD RESENDE, 2007), a maioria das medidas tradicionais de estabilidade podem ser enquadradas na metodologia de modelos mistos, assumindo os efeitos de genótipos como fixos e os efeitos de locais como aleatórios. O método REML é então aplicado na estimação de parâmetros, fato que é vantajoso devido à sua aplicabilidade para a situação de dados desbalanceados e de heterogeneidade de variâncias.

Atualmente, procedimentos de interpretação mais simples têm tido apelo para a análise de estabilidade e adaptabilidade. Neste sentido, medidas que incorporam ambos (estabilidade e adaptabilidade) em uma única estatística, tais quais os métodos de Annicchiarico (1992) e Lin e Binns (1988) e modificações, têm sido enfatizados (CRUZ e CARNEIRO, 2003). No contexto dos modelos mistos, um método para ordenamento de genótipos simultaneamente por seus valores genéticos (produtividade) e estabilidade, refere-se ao procedimento BLUP sob médias harmônicas (RESENDE, 2004). Neste caso, o vetor de dados (y) deve ser trabalhado como a recíproca dos dados observados, ou seja, $(1/y)$. Isto produz resultados que são função $(1/H)$ da média harmônica dos dados. Quanto menor for o desvio padrão do comportamento genotípico através dos locais, maior será a média harmônica de seus valores genotípicos através dos locais. Assim a seleção pelos maiores valores da média harmônica do valor genotípico (MHVG) implica simultaneamente na seleção para produtividade e estabilidade. Adicionalmente, o uso da transformação $1/y$ considera também a instabilidade dentro dos locais.

Em termos de adaptabilidade, uma medida simples e eficiente no contexto dos modelos mistos refere-se à performance relativa dos valores genotípicos (PRVG) através dos ambientes. Neste caso, os valores genotípicos preditos (ou os dados originais) são expressos como proporção da média geral de cada local e, posteriormente, obtém-se o valor médio desta proporção através dos locais. Genericamente, a performance relativa tem sido usada há um longo tempo (WRIGHT et al. 1966, apud RESENDE, 2007) em termos de dados fenotípicos, e constitui a base do método de Annicchiarico (1992 APUD RESENDE, 2007).

A seleção simultânea por produtividade, estabilidade e adaptabilidade, no contexto de modelos mistos, pode ser realizada pelo método da média harmônica da performance relativa do valor genotípico (MHPRVG) predito, proposto por Resende (2004).

De acordo com Resende (2007), este método permite selecionar simultaneamente pelos três atributos mencionados e apresenta as seguintes vantagens: (i) considera os efeitos genotípicos como aleatórios e portanto fornece a estabilidade e adaptabilidade genotípica e não fenotípica; (ii) permite lidar com desbalanceamento; (iii) permite lidar com delineamentos não ortogonais; (iv) permite lidar com heterogeneidade de variâncias; (v) permite considerar erros correlacionados dentro de locais; (vi) fornece valores genéticos já descontados (penalizados) da instabilidade; (vii) pode ser aplicado com qualquer número de ambientes; (viii) permite considerar a estabilidade e adaptabilidade na seleção de indivíduos dentro de progênies; (ix) não depende da estimação e

interpretação de outros parâmetros tais quais coeficientes de regressão; (x) elimina os ruídos da interação genótipos x ambientes pois considera a herdabilidade desses efeitos; (xi) gera resultados na própria grandeza ou escala do caráter avaliado; (xii) permite computar o ganho genético com a seleção pelos três atributos simultaneamente. Estes dois últimos fatores são bastante importantes. Outros métodos como o de Lin e Binns fornecem resultados que não são interpretados diretamente como valores genéticos e então não permitem computar o ganho genético no caráter composto pela produtividade, estabilidade e adaptabilidade. O método Annicchiarico depende, adicionalmente, de suposições de valores de α associados a $Z_{(1-\alpha)}$, que refere-se ao percentil da função distribuição normal padrão associado a determinado nível de significância α .

Oliveira et al. (2005), Sturion e Resende (2005) e Carbonell et al. (2007) avaliaram a produtividade, adaptabilidade e estabilidade pelo método da MHPRVG. Estes trabalhos mostraram que os métodos de Annicchiarico (1992), Lins e Binns (1988) e MHPRVG obtiveram, praticamente, o mesmo ranking de indivíduos selecionados ao serem comparados. Contudo, corroboraram que o método MHPRVG apresenta a vantagem de fornecer resultados na própria escala e medição do caráter avaliado, os quais podem ser interpretados diretamente como valores genéticos.

Trabalhos, como os de Pinto Jr. et al. (2006), Bastos et al. (2007), Zeni Neto et al. (2008), Maia et al. (2009), Verardi et al. (2009), Moraes et al. (2012), Pastrana Vargas et al. (2012), Rosado et al. (2012), Arantes et al. (2013), têm mostrado as vantagens de se utilizar a MHPRVG na seleção de plantas visando, simultaneamente, a produtividade, estabilidade e adaptabilidade em culturas perenes e semi-perenes, tais como: a seringueira, eucalipto, cajueiro, aroeira, *Acacia mangium*, cana-de-açúcar e etc.

A aplicação em culturas anuais também já foi realizada por Borges et al. (2010b) e Regitano Neto (2013) em arroz de terras altas e em milho por Mendes et al. (2012).

2.4.3. Norma de reação com covariáveis desconhecidas (SU et al., 2006)

A interação genótipo x ambiente (GxA) refere-se ao diferencial de respostas de diferentes genótipos através de uma variação de ambientes (KANG, 2004). Falconer and Mackay (1996) citam que a interação GxA é quando diferentes genótipos respondem diferentemente a mudanças em seus ambientes e chama-a, também, de diferenças na sensibilidade ambiental dos genótipos. Para análises de tal sensibilidade ambiental a

performance do genótipo (valor genotípico) é regredida sobre a performance média da população em cada ambiente.

Uma variação de fenótipos pode resultar de um genótipo quando este é exposto em diferentes ambientes. Este fenômeno é chamado de norma de reação, ou plasticidade fenotípica. Norma de reação representa a expressão da variabilidade de um fenótipo em indivíduos de um único genótipo (SAXTON, 2004). O valor genotípico estimado de um indivíduo, que é dividido numa dependente e independente parte ambiental, é também chamado norma de reação (FALCONER; MACKAY, 1996).

Para Scheiner (1993) a norma de reação representa a direção das respostas plásticas e a variabilidade entre os genótipos quanto ao grau de plasticidade, ou seja, descreve a expressão fenotípica pela genotípica como uma função do ambiente. A abordagem por norma de reação é usada quando fenótipos mudam gradualmente e continuamente sobre um gradiente ambiental (DE JONG, 1995).

Uma vez que os ambientes não são classificados em grupos, um método comum utilizado na norma de reação é a regressão da performance do genótipo (valor genotípico) em cada ambiente sobre o gradiente ambiental. Que seria, basicamente, o cálculo da média fenotípica de cada ambiente em que o genótipo foi avaliado e a utilização dessa média fenotípica ambiental como um índice ambiental e utilizá-lo como regressor na análise.

Van der Werf et al. (1998) utilizaram a análise de covariância na tentativa de captar a interação GxA realizando, num primeiro passo, o cálculo das médias fenotípicas e posteriormente, num segundo passo, a análise de regressão considerando estas médias fenotípicas ambientais como covariáveis na regressão.

A utilização de covariáveis para a captação da interação GxA praticamente começou com Kolmodin et al. (2002) que perceberam que apesar da desvantagem de incluir informação da variável dependente na variável explanatória, o método do cálculo de médias fenotípicas ambientais e utilização de tais médias como covariável (aqui chamado de procedimento de dois passos) provou-se utilizável. A maior vantagem deste procedimento foi que o ambiente pode ser pontuado como mais ou menos favorável. Calus and Veerkamp (2003) citaram que o uso de um parâmetro ambiental numa função de covariância tem a vantagem de que os ambientes são tratados como um contínuo, ao invés de um conjunto de grupos definidos arbitrariamente pelos dados.

Depois dessa introdução de análise por Kolmodin et al. (2002) houveram diversas pesquisas, tais como de Calus et al. (2002), Calus and Veerkamp (2003), Kolmodin and

Bijma (2004), Calus et al. (2004), Pollott and Greeff (2004), Calus et al. (2005), Pegolo et al. (2009), Corrêa et al. (2009), Corrêa et al. (2010) e Pegolo et al. (2011) que utilizaram da performance da média fenotípica do genótipo no ambiente e usam tal média no modelo como covariável desconhecida, ou seja, pelo procedimento de dois passos.

Calus et al. (2004) propuseram estimar os valores ambientais por um procedimento iterativo do qual o efeito ambiental estimado numa dada iteração substitui o valor da covariável na próxima. Utilizando simulação computacional, os autores observaram uma redução insignificante no viés das estimativas dos componentes de variância por essa abordagem quando comparado com o procedimento padrão de substituir covariáveis com médias fenotípicas. Eles sugeriram substituir os valores ambientais com estimativas dos efeitos de rebanhos obtidos de um grande número de animais por rebanho, ao invés de rebanho-anos, tendo o custo da perda de informação na Interação Genótipo x Ambiente.

Posteriormente ao início da utilização do procedimento de dois passos uma outra abordagem foi realizada e a utilização da média fenotípica do genótipo no ambiente como uma covariável no modelo já começou a ser provada que estavam ocorrendo perdas de informações na análise de interação GxA. Su et al. (2006) então afirmaram que é claramente insatisfatório incluir uma função dos dados como uma covariável no modelo amostral e que, além da atenuação da incerteza devido ao tratamento de médias fenotípicas como parâmetros conhecidos, pode-se imaginar situações que resultariam em representações enganosas dos valores ambientais.

Para superar essa limitação, Su et al. (2006) propuseram uma abordagem Bayesiana que infere conjuntamente os valores ambientais com todos os parâmetros desconhecidos na norma de reação, sem a necessidade de usar resultados de análise prévia (ou seja, um procedimento com um passo). O método foi testado utilizando-se dados simulados e conduziu para estimativas de parâmetros sem viés detectável e com um menor erro do quadrado médio do que os métodos convencionais de obtenção dos parâmetros ambientais.

Mais tarde foi demonstrado também por Shariati et al. (2007), Mattar et al. (2011) e Cardoso and Tempelman (2012) os benefícios de se estimar conjuntamente os parâmetros por esse tipo de abordagem em que consideram-se as covariáveis desconhecidas.

Shariati et al. (2007) estudaram a interação genótipo x ambiente em vacas dinamarquesas para diversas características de produção de leite a partir de modelos de norma de reação com covariáveis desconhecidas, utilizando-se a variância residual

homogênea e também variâncias residuais heterogêneas. Os autores concluíram que mesmo em um país pequeno como a Dinamarca a interação genótipo x ambiente pode ser detectada pelo modelo de norma de reação e sugeriu que a variância residual heterogênea deveria ser levada em consideração na análise de GxA. O uso da variância residual heterogênea na análise da norma de reação evitou o confundimento entre a variação do coeficiente de regressão genético e a variação dos efeitos ambientais dentro do sistema de produção.

3.REFERÊNCIAS BIBLIOGRÁFICAS

- ANNICCHIARICO, P. Cultivar adaptation and recommendation from alfafa trials in Northern Italy. **Journal of Genetics & Breeding**, Berlin, v. 46, n. 1, p. 269-278, 1992.
- ARANTES, F.C.; GONÇALVES, P.S.; SCALOPPI Junior, E.J.; MORAES, M.L.T.; RESENDE, M.D.V. Ganho genético com base no tamanho efetivo populacional de progênie de seringueira. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 12, p. 1419-1424, 2010.
- ARANTES, F.C.; SCALOPPI Jr., E.J.; GONÇALVES, P.S.; MORAES, M.L.T.; GONÇALVES, E.C.P.; RESENDE, M.D.V. Adaptability and stability in rubber tree progenies under different environmental conditions, **Pesquisa Florestal Brasileira**, Colombo, v. 33, n. 73, 2013.
- BAKER, R.J. Genotype-environment interaction in yield of wheat. **Can. J. Plant Sci.** 49, 743-751, 1969.
- BARTLETT, M. S. The use of transformations. **Biometrics**, 3, 39-57, 1947
- BASTOS, I. T.; BARBOSA, M. H. P.; RESENDE, M. D. V.; PETERNELLI, L.A.; SILVEIRA, L.C.I.; DONDA, L.R.; FORTUNATO, A. A.; COSTA, P. M. A.; FIGUEIREDO, I. C. R. Avaliação da interação genótipo x ambiente em cana-de-açúcar via modelos mistos. **Pesquisa Agropecuária Tropical**, v. 37, n. 4, 2007.
- BAUER, A.M.; REETZ, T.C.; LÉON, J. Estimation of breeding values of inbred lines using best linear unbiased prediction (BLUP) and genetic similarities. **Crop Science** v. 46, p. 2685 – 2691, 2006.
- BERNARDO, R. Best linear unbiased prediction maize single-cross performance. **Crop Science**, 36:50–56, 1996.
- BERNARDO, R. **Breeding for quantitative traits in plants**, 2nd edition, Minnesota, 390 p., 2010.
- BERNARDO, R. Prediction of maize single-cross performance using RFLPs and information from related hybrids. **Crop Science**, 34:20-25, 1994.

- BLASCO, A. The Bayesian controversy in animal breeding. **Journal of Animal Science**, 79, 2023–2046. 2001.
- 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, 2010a. DOI: 10.4025/actasciagron.v32i4.4837
- BORGES, V.; SOARES, A.A.; REIS, M.S.; RESENDE, M.D.V.; CORNÉLIO, V.M.O.; LEITE, N.A.; VIEIRA, A.R.; Desempenho genotípico de linhagens de arroz de terras altas utilizando metodologia de modelos mistos. **Bragantia**, Campinas, v. 69, n. 4, p. 833-841, 2010b.
- BREESE, E.L. The measurement and significance of genotype-environment interactions in grasses. **Heredity**, 24, 27-44, 1969.
- BUCIO ALANIS, L. Environmental and genotype-environment components of variability I- Inbred Lines. **Heredity**, 31, 387-397, 1966.
- BUCIO ALANIS, L.; HILL, J. Environmental and genotype environmental components of variability II. Heterozygotes. **Heredity**, 21, 399-405, 1966.
- CALUS, M. P. L.; BIJMA, P.; VEERKAMP, R. F. Effects of data structure on the estimation of covariance functions to describe genotype by environment interaction in a reaction norm model. **Genet. Sel. Evol.** v. 36, p. 489–507, 2004.
- CALUS, M. P. L.; GROEN, A. F.; DE JONG, G. Genotype ×environment interaction for protein yield in Dutch dairy cattle as quantified by different models. **Journal of Dairy Science**. v. 85, p. 3115–3123, 2002
- CALUS, M. P. L.; VEERKAMP, R. F. Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. **Journal of Dairy Science**, v. 86, p. 3756–3764, 2003.
- CALUS, M.P.L.; CARRICK, M.J.; VEERKAMP, R.F.; GODDARD, M.E. Estimation of genetic parameters for milk fat depression in dairy cattle. **Journal of Dairy Science**, Champaign, v. 88, p. 1166-1177, 2005.
- CARBONELL, S. A. M.; CHIORATO, A. F.; RESENDE, M. D. V.; DIAS, L. A. S.; BERALDO, A. L. A.; PERINA, E. F. Estabilidade de cultivares e linhagens de feijoeiro em diferentes ambientes no estado de São Paulo. **Bragantia**, Campinas, v. 66, n. 2, p.193-201, 2007.
- CARDOSO, F.F.; TEMPELMAN, R.J. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. **Journal of Animal Science**, v. 90, p. 2130-2141, 2012.
- CASELLA, G.; GEORGE, E.I. Explaining the Gibbs Sampler. **The American Statistician**, v. 46, n. 3, p. 167-174, 1992.
- CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Caracterização da interação genótipo-ambiente e comparação entre modelos para ajuste do ganho pós-

demama de bovinos Devon via normas de reação. **Revista Brasileira de Zootecnia**, v.38, n.8, p.1468-1477, 2009.

CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Genetic evaluation of Devon Cattle using a reaction norms model. **Revista Brasileira de Zootecnia**, v.39, n.1, p.128-133, 2010.

CROSSA, J.; BURGUENO, J.; CORNELIUS, P.L.; MCLAREN, G.; TRETHOWAN, R.; KRISHNAMACHARI, A. Modeling genotype x environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. **Crop Sci** 46:1722–1733, 2006.

CRUZ, C.D.; CARNEIRO, P.C.S.; **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2 ed., v. 2, 2003. 585p.

DAHER, R. F.; PEREIRA, M. G.; AMARAL JR., A. T.; PEREIRA, A. V.; LÉDO, F. J. S.; DAROS, M. Estabilidade da produção forrageira em clones de capim-elefante (*Pennisetum purpureum* Schum.). **Ciência Agrotecnologia**, v. 27, n. 4, p. 788-797, 2003.

DANIELS, J.; ROACH, B.T. Taxonomy and evolution. p.7-84. *In* Heinz, D.J. (ed.) **Sugarcane improvement through breeding**. Elsevier Press, Amsterdam. 1987.

DE JONG, G. Phenotypic plasticity as a product of selection in a variable environment. **The American Naturalist**, Chicago-IL, V. 145, 493-512, 1995.

EAGLES, H.A.; MOODY, D.B. Using unbalanced data from a barley breeding program to estimate gene effects: Ha2, Ha4, and sdw1 genes. **Australian Journal of Agricultural Research**, v. 55, p. 379–387, 2004.

EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, n. 1, p. 36-40, 1966.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**, 4 th ed. Longman Group, Essex, 464 p. 1996.

FARIAS NETO, J.T.; LINS, P.M.P.; RESENDE, M.D.V.; MULLER, A.A. Seleção genética em progênies híbridas de coqueiro. **Revista Brasileira de Fruticultura**, v. 31, p. 190-196, 2009.

FINLAY, K.W.; WILKINSON, G.N. The analysis of adaptation in a plant breeding 24 programme. **Aust. J. Agri. Res.**, 14, 742-754, 1963.

FISHER, R.A.; MACKENZIE, W.A. Studies in crop variation II. The response of different potato varieties. **J. Agric. Sci.**, Camb., 13, 311-320, 1923.

FLACHENECKER, C.; FRISCH, M.; FALKE, K.C.; MELCHINGER, A.E. Modified full-sib selection and best linear unbiased prediction of progeny performance in a European F2 maize population. **Plant Breed** 125:248–253, 2006a.

FLACHENECKER, C.; FRISCH, M.; FALKE, K.C.; MELCHINGER, A.E. Trends in population parameters and best linear unbiased prediction of progeny performance in a European F2 maize population under modified recurrent full-sib

selection. **Theor Appl Genet** 112:483–491, 2006b.

FOX, P.N.; CROSSA, J.; ROMAGOSA, I. Multi-environment testing and genotype x environment interaction. In: KEMPTON, R.A.; FOX, P.N. (Ed.). **Statistical methods for plant variety evaluation**. New York: Chapman and Hall, Cap. 8, p. 117-138, 1997.

FREEMAN, A.E; HENDERSON, C.R. Contributions to the dairy industry. **Journal of Dairy Science**. 74: 4045-4051, 1991.

FREEMAN, G.H.; PERKINS, J.M. Environmental and genotype-environmental components of variability VIII. Relation between genotype grown in different environments and measurement of these environments. **Heredity**, 27, 15-23, 1971.

GAMERMAN, D. **Markov chain Monte Carlo**: Stochastic simulation for bayesian inference. CRC Press, 1997. 230 p.

GELMAN, A.; RUBIN, D.B. Inference from iterative simulation using multiple sequence. **Statistical Science**, Hayward, v.7, n.4, p.457-511, 1992.

GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distribution and Bayesian restoration of images. **IEE Transactions on Pattern Analysis and Machine Intelligence**, v. 6, p. 721–741, 1984.

GEWEKE, J. **Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments**. In, Bayesian Statistics 4, J. M. Bernardo, J. O. Berger, A. P. Dawid,; A. F. M. Smith (eds.), pp. 169-193. Oxford University Press. 1992.

GIANOLA, D.; FERNANDO, R.L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v. 63, p. 217-244, 1986.

GIANOLA, D.; FERNANDO, R.L.; IM, S.; FOULLEY, J.L. Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. **Genome**, Ottawa, v. 31, p. 768-777, 1989.

HARDWICK, R.C.; WOOD, J.T. Regression methods for studying genotype- environment interactions. **Heredity**, 28, 209-222, 1972.

HARVILLE, D. A. "Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects." **Annals of Statistics** 2:384–395, 1976.

HARVILLE, D.A. Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, Washington, v. 72, p. 320-328, 1977.

HENDERSON, C. R. Selection Index and Expected Genetic Advance. In: **Statistical Genetics and Plant Breeding**, ed. W. D. Hanson and H. F. Robinson, 141–163. Washington, DC: National Academy of Sciences and National Research Council, Publication No. 982, 1963.

HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection

model. **Biometrics** 31: 423-447, 1975.

HENDERSON, C.R.; KEMPTHORNE, O.; SEARLE, S.R.; VON KROSIGK, C.M. The estimation of environmental and genetic trends from records subject to culling. **Biometrics** 15: 192-218, 1959.

HOGARTH, D.M. Quantitative inheritance studies in sugarcane. I. Estimation of competition and violation of genetic assumptions on estimation of genetic variance components. **Australian Journal of Agricultural Research**, v. 28, p. 257-268. 1977.

HOGG, R.V.; CRAIG, A.T. **Introduction to mathematical statistics**, 2^a ed., New York, Macmillan Co., 1965. 388 p.

KANG, M.S. 2004. **Breeding: Genotype by environment interaction**. In Encyclopedia of Plant and Crop Science, ed. R.M. Goodman. New York: Marcel Dekker.

KENNEDY, B. W.; SORENSEN, D.A. Properties of mixed model methods for prediction of genetic merit under different genetic models in selected and unselected populations. In: WEIR, B.; GOODMAN, M.M.; NAMKOONG, G. (Ed.). **Second International Conference on Quantitative Genetics**, Raleigh, 1988. Raleigh: North Carolina State University, 1988. p. 91-103.

KOLMODIN, R., E. STRANDBERG, MADSEN, P., JENSEN, J. ; JORJANI, H. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. **Acta Agriculturae, Scandinavica**, Section A: Animal. Science. 52:11–24, 2002.

KOLMODIN, R.; BIJMA, P. Response to mass selection when the genotype by environment interaction is modeled as a linear reaction norm. **Genet. Sel. Evol.** v. 36, p. 435-454, 2004.

LIN, C.S.; BINNS, M.R. A method of analysing cultivars x location x year experiments: a new stability parameter. **Theoretical Applied Genetics**, Berlim, v.76, p.425-430, 1988.

LITTELL, RAMON C., GEORGE A. MILLIKEN, WALTER W. STROUP, RUSSELL D. WOLFINGER, OLIVER SCHABENBERGER. **SAS® for Mixed Models**, Second Edition. Cary, NC: SAS Institute Inc, 2006.

LYNCH, M.; WALSH, B. **Genetics and analysis of quantitative traits**. Sunderland: Sinauer Associates, Inc., 1998, 980p.

MAIA, M.C.C.; RESENDE, M.D.V.; PAIVA, J.R. de; CAVALCANTI, J.J.V.; BARROS, L. de M.B. Seleção simultânea para produção, adaptabilidade e estabilidade genotípicas em clones de cajueiro, via modelos mistos. **Pesquisa Agropecuária Tropical**, v.39, p.43-50, 2009.

MATSUOKA, S.; GARCIA, A.A.F.; ARIZONO, H. Melhoramento da cana-de-açúcar. In: BOREM, A. (Ed.). **Melhoramento de espécies cultivadas**. Viçosa: Editora UFV, 2005. p. 225-274.

MATTAR, M.; SILVA, L.O.C.; ALENCAR, M.M. ; CARDOSO, F.F.; Genotype by

environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. **Journal of Animal Science**. v. 89, p. 2349-2355, 2011.

MENDES, F.F.; GUIMARÃES, L.J.M.; SOUZA, J.C.; GUIMARÃES, P.E.O.; PACHECO, C.A.P.; MACHADO, J.R.A.; MEIRELLES, W.F.; SILVA, A.R.; PARENTONI, S.N. Adaptability and stability of maize varieties using mixed model methodology. **Crop Breeding and Applied Biotechnology**, v. 12, p. 111-117, 2012.

MEYER, K. Estimation of genetic parameters. In: HILL, W.G.; MACKAY, T.F.C. **Evolution and animal breeding**, Wallingford: CAB International, 1989. p. 161-167.

MI, X.; WEGENAST T.; UTZ, H.F.; DHILLON, B.S.; MELCHINGER, A.E. Best linear unbiased prediction and optimum allocation of test resources in maize breeding with doubled haploids. **Theoretical and Applied Genetics** 123:1-10, 2011.

MILLER, P.A.; WILLIAMS, J.C.; ROBINSON, H.F. Variety x environment interactions in cotton variety tests and their implications on testing methods. **Agron. J.**, 51, 132-134, 1959.

MISSIO, R.F.; DIAS, L.A.S.; MORAES, M.L.T.; RESENDE, M.D.V. Selection of *Pinus caribaea* var. *bahamensis* progênies based on the predicted genetic value. **Crop Breeding and Applied Biotechnology**, v. 4, p. 399-407, 2004.

MISSIO, R.F.; SILVA, A.M.; DIAS, L.A.S.; MORAES, M.L.T.; RESENDE, M.D.V. Estimates of genetic parameters and prediction of additive genetic values in *Pinus kesya* progenies. **Crop Breeding and Applied Biotechnology**, v. 5, p. 394-401, 2005.

MORAES, M.A.; VALÉRIO FILHO, W.V.; RESENDE, M.D.V.; SILVA, A.M.; MANOEL, R.O.; FREITAS, M.L.M.; MORAES, M.L.T.; SEBBENN, A.M. Produtividade, estabilidade e adaptabilidade em progênies de *Myracrodruon urundeuva* F.F. & M.F. Allemão - Anacardiaceae. **Scientia Forestalis**, Piracicaba, v. 40, n. 93, p. 069-076, 2012.

MORELY, J.R.; MATHER, K. Interaction of genotype and environment in continuous variation II. Analysis. **Biometric.**, 14, 481-488, 1958.

NUNES, J.A.R.; RAMALHO, M.A.P.; FERREIRA, D.F. Inclusion of genetic relationship information in the pedigree selection method using mixed models. **Genet Mol Biol** 31:73-78, 2008.

OLIVEIRA, R.; RESENDE, M. D. V. ; DAROS, E. ; BESPALHOK FILHO, J. C.; ZAMBON, J. L.; IDO, O.; KOEHLER, Henrique . Genotypic evaluation and selection of sugarcane clones in three environments in the state of Paraná. **Crop breeding and applied biotechnology**, Viçosa, v. 5, n. 4, p. 426-434, 2005.

PAIVA, J.R.; RESENDE, M.D.V.; CORDEIRO, E.R. Índice multifeitos e estimativas de parâmetros genéticos aplicados ao melhoramento da acerola. **Pesquisa Agropecuária Brasileira**, v. 37, p. 799 – 807, 2002.

PANTER, D.M.; ALLEN, F.L. Using best linear unbiased predictions to enhance breeding for yield in soybean: I Choosing parents. **Crop Science**, 35:397-405, 1995.

PASTRANA VARGAS, I.J.; ESPÍLIA, C.M.; MURILLO, G.O. Evaluación del potencial de

mejoramiento genético en el crecimiento en altura de *Acacia mangium* Willd. **Acta Agronômica**, v. 61, n. 2, 2012.

PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika** 58: 545-554, 1971.

PEDROZO, C.A.; BENITES, F.R.G.; BARBOSA, M.H.P.; RESENDE, M.D.V.; SILVA, F.L. Eficiência de índices de seleção utilizando a metodologia REML/BLUP no melhoramento da cana-de-açúcar. **Scientia Agraria**, v. 10, p. 31-36, 2009.

PEGOLO, N.T.; ALBUQUERQUE, L.G.; LÔBO, R.B.; OLIVEIRA, H.N. Effects of sex and age on genotype x environment interaction for beef cattle weight studied using reaction norm models. **Journal of Animal Science**, v. 89, p. 3410-3425, 2011.

PEGOLO, N.T.; OLIVEIRA, H.N.; ALBUQUERQUE, L.G.; BEZERRA, L.A.F.; LÔBO, R.B. Genotype by environment interaction for 450-day weight of Nelore cattle analyzed by reaction norm models. **Genetic and Molecular Biology**, v. 32, n.2, p. 281-287, 2009.

PERKINS, J. M. The principal component analysis of interactions and physical measures of the environment. *Heredity*, 29, 51-70, 1972.

PERKINS, J. M.; JINKS, J. L. Environmental and GE components of variability, IV. Non-linear interactions for multiple inbred lines. **Heredity**, 23, 525-535, 1968b.

PERKINS, J. M.; JINKS, J. L. Environmental and genotype environmental components of variability, III. Multiple lines and crosses. **Heredity**, 23, 339-356. 1968a.

PIEPHO HP, MOHRING J, MELCHINGER AE, BUCHSE A. BLUP for phenotypic selection in plant breeding and variety testing. **Euphytica** 161:209–228, 2008.

PIEPHO, H.P. Stability analysis using SAS, **Agronomy Journal**, v. 91, p. 154-160, 1999.

PINTO Jr, J.E.; STURION, J.A.; RESENDE, M.D.V.; RONZELLI Jr, P.R.; Avaliação simultânea de produtividade, adaptabilidade e estabilidade genotípica de *Eucalyptus grandis* em distintos ambientes do Estado de São Paulo. **Boletim de Pesquisa Florestal**, Colombo, n. 53, p. 79-108, 2006.

POLLOTT, G.E.; GREEFF, J.C. Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. **Journal of Animal Science**, Champaign, v. 82, p. 2840-2851, 2004.

RAFTERY, A.E.; LEWIS, S.M. Comment: One long run with diagnostics: Implementation strategies for Markov Chain Monte Carlo. **Stat. Sci.**, v. 7, p. 493–497, 1992a.

RAFTERY, A.E.; LEWIS, S.M. **How many iterations in the Gibbs sampler?** In: Bayesian Statistics 4, J. M. Bernardo, J. O. Berger, A. P. Dawid, ; A. F. M. Smith (eds.), p. 763–773. Oxford University Press. 1992b.

REGITANO NETO, A.; RAMOS Jr, E.U.; GALLO, P.B.; FREITAS, J.G.; AZZINI, L.E. Comportamento de genótipos de arroz de terras altas no estado de São Paulo. **Revista Ciência Agronômica**, v. 44, n. 3, p. 512-519, 2013.

- REIS, A.J.D.; CHAVES, L.J.; DUARTE, J.B.; BRASIL, E.M. Prediction of hybrid means from a partial circulant diallel table using the ordinary least square and the mixed model methods. **Genet Mol Biol** 28:314–320, 2005.
- RESENDE, M. D. V. **Métodos Estatísticos Ótimos na Análise de Experimentos de Campo**. Colombo: Embrapa Florestas, Ed. 1, 2004, 57 p. (Documentos, 100).
- RESENDE, M.D.V. de. Avanços da genética biométrica florestal. In: BANDEL, G.; VELLO, N.A.; MIRANDA FILHO, J.B. (Ed.). **Encontro sobre temas de genética e melhoramento: genética biométrica vegetal**. Anais. Piracicaba: Esalq/Usp, 1997. p. 20-46.
- RESENDE, M.D.V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília: EMBRAPA Informação Tecnológica, 2002. 975p.
- RESENDE, M.D.V. **Inferência Bayesiana e simulação estocástica (amostragem de Gibbs) na estimação de componentes de variância e de valores genéticos em plantas perenes**. Colombo: Embrapa Florestas, 2000. 68p. (Documentos, 46).
- RESENDE, M.D.V. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas. 2007. 561p.
- ROBERTSON, A. Weighting in the estimation of variance components in the unbalanced single classification. **Biometrics** 18 : 413–417, 1962.
- 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**, Brasília, v. 47, n. 7, p. 964-971, 2012.
- ROW, P.R.; ANDREW, R.A. Phenotypic stability for a systematic series of corn genotypes. *Crop. Sci.*, 4, 553-567, 1964.
- SAXTON, A. Ed. 2004. **Genetic Analysis of Complex Traits Using SAS®**. Cary, NC: SAS Institute Inc.
- SCHAEFFER, L.R.; HENDERSON, C.R. Contributions to predicting genetic merit. *Journal of Dairy Science*. 74: 4052-4066, 1991.
- SCHEINER, S.M. Genetics and evolution of phenotypic plasticity. **Annual Review of Ecologic Systems**, Palo Alto, V. 24, p. 35-68, 1993.
- SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. Variance components. 2.ed. New York : J. Willey, 1992. 528p.
- SEBOLAI, B.; PEDERSEN, J.F.; MARX, D.B.; BOYKIN, D.L. Effect of control plot density, control plot arrangement, and assumption of random or fixed effects on nonreplicated experiments for germplasm screening using spatial models. **Crop Science** 45:1978–1984, 2005.
- SHARIATI, M.M.; SU, G.; MADSEN, P.; SORENSEN, D. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. **Journal**

of **Dairy Science**, v. 90, p. 5759-5766, 2007.

SHUKLA, G.K. Some statistical aspects of partitioning genotype-environmental components of variability. **Heredity**, 29, 237-245, 1972.

SHUKLA, G.K. Use of regression approach in the analysis of genotype- environmental interaction. **J. Ind. Soc. Agr. Stat.**, 35(2), 129-138, 1983.

SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian and MCMC methods in quantitative genetics**, New York, Springer, 2002, 740 p.

SORENSEN, D.A. **Gibbs sampling in quantitative genetics**. Tjele: Danish Institute of Animal Science, Department of Breeding and Genetics. Intern Report n. 82. 1996. 186 p.

SOUZA, V.A.B.; BYRNE, D.H.; TAYLOR, J.F. Predicted breeding values for nine plant and fruit characteristics of 28 peach genotypes. **Journal of the American Society for Horticultural Science**, v. 125, p. 460–465, 2000.

STURION, J. A.; RESENDE, M.D.V. Seleção de progênies de erva-mate (*Ilex paraguariensis* St. Hil.) para produtividade, estabilidade e adaptabilidade temporal de massa foliar. **Boletim de pesquisa florestal**, Colombo, n. 50, p. 37-51, 2005.

SU, G.; MADSEN, P.; LUND, M. S.; D, SORENSEN; KORSGAARD, I.R.; JENSEN, J. Bayesian analysis of the linear reaction norm model with unknown covariates. **Journal of Animal Science**, v. 84, p. 1651–1657, 2006.

TUKEY, J.W. One degree of freedom for non-additivity. **Biometrics**, 5, 232-242, 1949.

UNICA, União da Indústria de cana-de-açúcar, Fonte: www.unicadata.com.br, Acesso: 05 set. 2013.

VAN DER WERF, J. H. J., M. E. GODDARD, ; K. MEYER. 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. **Journal of Dairy Science** 81:3300–3308.

VENCOVSKY, R.; BARRIGA, P.; **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992, 486p.

VERARDI, C.K.; RESENDE, M.D.V.; COSTA, R.B.; GONÇALVES, P.S. Adaptabilidade e estabilidade da produção de borracha e seleção em progênies de seringueira. **Pesquisa Agropecuária Brasileira**, v. 44, n. 10, p. 1277-1282, 2009.

VIANA, J.M.S.; ALMEIDA, I.F.; RESENDE, M.D.V.; FARIA, V.R.; SILVA, F.F. BLUP for genetic evaluation of plants in non-inbred families of annual crops. **Euphytica**, v. 174, p. 31-39, 2010a.

VIANA, J.M.S.; SOBREIRA, F.M.; RESENDE, M.D.V.; FARIA, V.R. Multi-trait BLUP in half-sib selection of annual crops. **Plant Breeding**, v. 129, p. 599-604, 2010b.

WRIGHT, J.W.; PAULEY, S.S.; POLK, R.B.; JOKELA, J.J. Performance of Scotch pine varieties in North Central Region, **Silvae Genetica**, v. 15, 0. 101-110, 1966.

- YAN, W.; KANG, M.S. 2003. GGE biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists. Boca Raton, FL: CRC Press.
- YATES, F.; COCHRAN, W.G. The analysis of group of experiments. J. Agric. Sci., Cambridge, 28, 556-580, 1938.
- ZENI NETO, H.; OLIVEIRA, R.A. de; DAROS, E.; BESPALHOK FILHO, J.C.; ZAMBON, J.L.C.; IDO, O.T.; WEBER, H. Seleção para produtividade, estabilidade e adaptabilidade de clones de cana-de-açúcar em três ambientes no Estado do Paraná via modelos mistos. **Scientia Agraria**, v.9, p.425-430, 2008.

CAPÍTULO 2 - MATERIAL E MÉTODOS

1. MATERIAL

O conjunto de dados foi coletado de ensaios do Procana IAC, instalados em 2007, referente a quatro cortes de 2008 a 2011 (cana planta, primeira, segunda e terceira soca). Avaliou-se 16 genótipos num total de 14 usinas (ambientes). Desse total, 13 usinas foram avaliadas no primeiro corte, 12 usinas no segundo corte, 10 no terceiro e sete no quarto corte. As usinas encontram-se descritas no software CAIANA sob domínio do IAC - Ribeirão Preto, Estado de São Paulo, Brasil e correspondem a quatro estados da região centro-sul do Brasil. Os municípios em que se encontram as usinas situam-se nos Estados de Goiás (Goianésia, Goiatuba, Jandaia e Turvelândia), Mato Grosso (Nova Olímpia), Minas Gerais (Delta) e São Paulo (Araraquara, Catanduva, Guaíra, Guariba, Iracemápolis, São João da Boa Vista, São Joaquim da Barra e Severínia). Os atributos mensurados foram a Tonelada de Colmos por Hectare (TCH) e porcentagem aparente de sacarose (POL%) transformado em somente um atributo, Tonelada de Pol por Hectare (TPH), pela fórmula: $TPH = (TCH * POL\%) / 100$. O TPH foi o atributo submetido à análise. O delineamento experimental aplicado para cada ambiente foi de blocos completos com três repetições. A parcela foi composta de cinco linhas de dez metros e o espaçamento entre linhas foi de 1,5 m. O TPH foi mensurado em nove genótipos testes: IACSP952289, IACSP962008, IACSP962073, IACSP963048, IACSP963076, IACSP972084, IACSP985024, IACSP985046 e IACSP986202; e em sete variedades padrões: RB72454, RB867515, RB92579, SP813250, CTC2, CTC4 e IACSP933046. A escolha das variedades padrões foi em função da adaptabilidade e estabilidade demonstrada por época e região do ambiente onde elas foram avaliadas previamente assim como a representatividade da área de cultivo da cultura.

2. MÉTODOS

2.1. Regressão Linear Fenotípica Conjunta (EBERHART; RUSSELL, 1966)

Eberhart and Russell (1966) expandiram o modelo proposto por Finlay e Wilkinson (1963), sob o aspecto de que tanto os coeficientes de regressão dos valores fenotípicos de cada genótipo em relação ao índice ambiental quanto os desvios desta regressão proporcionam estimativas de parâmetros de estabilidade e adaptabilidade. Assim, pela metodologia destes autores, é considerado que um genótipo com coeficientes de regressão superior a 1,0 tem comportamento consistentemente melhor em ambientes favoráveis. A magnitude e a significância dos desvios de regressão dão uma estimativa da previsibilidade do material genético (CRUZ; REGAZZI, 2001).

Eberhart and Russell (1966) consideram como genótipo ideal aquele que apresenta alta produção média, coeficiente de regressão igual a 1,0 e desvios de regressão não significativos.

Metodologias de análises de adaptabilidade e estabilidade, baseadas em análise de regressão linear simples têm sido comumente utilizadas, entretanto os conceitos de mais fácil entendimento têm sido os seguintes:

Adaptabilidade: refere-se à capacidade de os genótipos aproveitarem vantajosamente o estímulo do ambiente. Quanto à adaptabilidade, classificam-se em:

- a) Genótipos com adaptabilidade geral ou ampla: são aqueles com coeficiente de regressão igual a 1;
- b) Genótipos com alta adaptabilidade a ambientes favoráveis: são aqueles com coeficiente de regressão maior que 1.
- c) Genótipos com adaptabilidade a ambientes desfavoráveis: são aqueles com coeficiente de regressão menor que 0.

Estabilidade: refere-se à capacidade de os genótipos mostrarem um comportamento previsível em função do estímulo do ambiente. É avaliada pelo componente atribuído aos desvios de regressão, sendo verificados os seguintes tipos de genótipos:

- a) Genótipos com estabilidade ou previsibilidade alta: são aqueles com desvios de regressão igual 0 (desvios de regressão não significativos ao nível de 5% de probabilidade).
- b) Genótipos com estabilidade ou previsibilidade baixa: são aqueles com desvios de regressão diferente de 0 (significativos ao nível de 5% de probabilidade).

Para estimação conjunta dos parâmetros do modelo de Eberhart and Russell (1966) utilizou-se o modelo

$$y_{ijk} = \mu + g_i + b_i I_j + d_i + \text{bloco}_k(I_j) + e_{ijk}$$

Onde:

y_{ijk} : observação fenotípica do genótipo i no ambiente j no bloco k ;

$\mu + g_i$: média observada do genótipo i ;

b_i : coeficiente de regressão linear do genótipo i em função dos índices ambientais (I_j);

I_j : Índice ambiental, definido como a diferença entre as médias dos genótipos no ambiente j e a média geral;

d_i : Desvios de regressão linear do genótipo i em função dos I_j

$\text{bloco}_k(I_j)$: Efeito do bloco k dentro do ambiente j ;

e_{ijk} : Efeito residual dentro dos ambientes.

Esquema da Análise de Variância Conjunta para I genótipos e J ambientes

Fonte de Variação	GL	Objetivo
Genótipo	$I-1$	Testar igualdade de médias de genótipos
Ambiente linear genótipo 1	1	Testar efeito linear para genótipo 1
Desvio ambiente linear gen. 1	$J-2$	Testar desvios do efeito linear para genótipo 1
...		
Ambiente linear genótipo I	1	Testar efeito linear para genótipo I
Desvio ambiente linear gen. I	$J-2$	Testar desvios do efeito linear para genótipo I
Blocos dentro de ambiente	$J(K-1)$	Ajustar seus efeitos
Resíduos dentro de ambiente	$J(I-1)(K-1)$	Estimar variância ambiental dentro de ambientes
Total	$I*J*K-1$	

GL: grau de liberdade.

As análises estatísticas para esta metodologia foram realizadas utilizando-se o software SAS, procedimento PROC GLM (SAS INSTITUTE INC., 2010).

2.2. Média Harmônica da Performance Relativa do Valor Genotípico (RESENDE, 2004)

Para a avaliação dos genótipos no delineamento de blocos completos, em vários

ambientes e uma observação por parcela foi utilizado o seguinte modelo estatístico:

$$y = \mathbf{Xr} + \mathbf{Tb} + \mathbf{Zg} + \mathbf{Wi} + e$$

Em que y é o vetor das observações, r é o vetor dos efeitos de locais (assumidos como fixos) somados à média geral, b é o efeito de bloco dentro de locais (assumidos como aleatórios), g é o vetor dos efeitos genotípicos (assumidos como aleatórios), i é vetor dos efeitos da interação genótipo x ambiente (aleatórios) e e é o vetor de erros ou resíduos (aleatórios). As letras maiúsculas representam as matrizes de incidência para os referidos efeitos.

Distribuição e estrutura das médias e variâncias:

$$y|r, V \sim N(\mathbf{Xr}, V)$$

$$b|\sigma_b^2 \sim N(0, \mathbf{I}\sigma_b^2)$$

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

$$i|\sigma_i^2 \sim N(0, \mathbf{I}\sigma_i^2)$$

$$e|\sigma_e^2 \sim N(0, \mathbf{I}\sigma_e^2)$$

$$\text{Cov}(b, g') = 0;$$

$$\text{Cov}(b, i') = 0;$$

$$\text{Cov}(b, e') = 0;$$

$$\text{Cov}(g, i') = 0;$$

$$\text{Cov}(g, e') = 0;$$

$$\text{Cov}(i, e') = 0;$$

ou seja:

$$E \begin{bmatrix} y \\ b \\ g \\ i \\ e \end{bmatrix} = \begin{bmatrix} \mathbf{Xr} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \quad \text{Var} \begin{bmatrix} y \\ b \\ g \\ i \\ e \end{bmatrix} = \begin{bmatrix} V & \mathbf{TP} & \mathbf{ZG} & \mathbf{WC} & \mathbf{R} \\ \mathbf{T}' & \mathbf{P} & 0 & 0 & 0 \\ \mathbf{Z}' & 0 & \mathbf{G} & 0 & 0 \\ \mathbf{W}' & 0 & 0 & \mathbf{C} & 0 \\ \mathbf{R} & 0 & 0 & 0 & \mathbf{R} \end{bmatrix}, \text{ em que:}$$

$$\mathbf{P} = \mathbf{I}\sigma_b^2; \quad \mathbf{G} = \mathbf{I}\sigma_g^2; \quad \mathbf{C} = \mathbf{I}\sigma_i^2; \quad \mathbf{R} = \mathbf{I}\sigma_e^2;$$

$$\mathbf{V} = \mathbf{TPT}' + \mathbf{ZGZ}' + \mathbf{WCW}' + \mathbf{R};$$

O ajuste do modelo é obtido a partir das equações do modelo misto:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{T} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{W} \\ \mathbf{T}'\mathbf{X} & \mathbf{T}'\mathbf{T} + \mathbf{I}\lambda_1 & \mathbf{T}'\mathbf{Z} & \mathbf{T}'\mathbf{W} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{T} & \mathbf{Z}'\mathbf{Z} + \mathbf{G}^{-1}\lambda_2 & \mathbf{Z}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{T} & \mathbf{W}'\mathbf{Z} & \mathbf{W}'\mathbf{W} + \mathbf{I}\lambda_3 \end{bmatrix} \begin{bmatrix} \hat{r} \\ \hat{b} \\ \hat{g} \\ \hat{i} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'y \\ \mathbf{T}'y \\ \mathbf{Z}'y \\ \mathbf{W}'y \end{bmatrix}, \text{ em que:}$$

$$\lambda_1 = \frac{\sigma_e^2}{\sigma_b^2} = \frac{1 - \rho}{b^2}; \quad \lambda_2 = \frac{\sigma_e^2}{\sigma_g^2} = \frac{1 - \rho}{h_g^2}; \quad \lambda_3 = \frac{\sigma_e^2}{\sigma_i^2} = \frac{1 - \rho}{i^2};$$

Onde:

$\rho = \frac{\hat{\sigma}_b^2 + \hat{\sigma}_g^2 + \hat{\sigma}_i^2}{\hat{\sigma}_b^2 + \hat{\sigma}_g^2 + \hat{\sigma}_i^2 + \hat{\sigma}_e^2}$: Repetibilidade individual dentro de ambiente;

$b^2 = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_b^2 + \hat{\sigma}_g^2 + \hat{\sigma}_i^2 + \hat{\sigma}_e^2}$: coeficiente de determinação do efeito de bloco;

$h_g^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_b^2 + \hat{\sigma}_i^2 + \hat{\sigma}_e^2}$: herdabilidade individual no sentido amplo dentro do bloco;

$se = \sqrt{\frac{\hat{h}_g^2}{n} \frac{GL_g}{n}} / 2$: erro padrão de h_g^2 ; GL_g : grau de liberdade dos genótipos; n: número de observações utilizadas para as estimativas;

$IC(\hat{h}_g^2)_{0,95} = \left[\hat{h}_g^2 - \sqrt{\frac{\hat{h}_g^2}{n} \frac{GL_g}{n}} \leq \hat{h}_g^2 \leq \hat{h}_g^2 + \sqrt{\frac{\hat{h}_g^2}{n} \frac{GL_g}{n}} \right]$: Intervalo de confiança para as estimativas dos coeficientes de herdabilidade individual no sentido amplo dentro do bloco;

$h_{mg}^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + (\hat{\sigma}_b^2/b) + (\hat{\sigma}_i^2/a) + (\hat{\sigma}_e^2/a*b)}$: Herdabilidade média genotípica; onde, b: número de blocos (3); a: número de ambientes (13:corte 1; 12:corte 2; 10: corte 3; 7: corte 4);

$\hat{r}_{gg} = \sqrt{h_{mg}^2}$: Acurácia de seleção de genótipos;

$i^2 = \frac{\hat{\sigma}_i^2}{\hat{\sigma}_b^2 + \hat{\sigma}_g^2 + \hat{\sigma}_i^2 + \hat{\sigma}_e^2}$: Coeficiente de determinação do efeito ambiental dentro do bloco;

$\hat{\sigma}_b^2$: Variância do efeito de blocos;

$\hat{\sigma}_g^2$: Variância genotípica entre os genótipos;

$\hat{\sigma}_i^2$: Variância da interação genótipo X ambiente;

$\hat{\sigma}_e^2$: Variância residual

$\hat{r}_{gloc} = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_i^2}$: Correlação genotípica entre as performances através dos ambientes;

$CV_g\% = \frac{\sqrt{\hat{\sigma}_g^2}}{\mu} \times 100$: Coeficiente de variação genotípica;

$CV_e\% = \frac{\sqrt{\hat{\sigma}_e^2}}{\mu} \times 100$: Coeficiente de variação residual;

μ : média geral

Os estimadores iterativos dos componentes de variância, por REML (Restricted

Maximum Likelihood), via algoritmo EM, são:

$$\hat{\sigma}_e^2 = [y'y - \hat{r}'X'y - \hat{b}'T'y - \hat{g}'Z'y - \hat{i}'W'y]/[N - r(x)]$$

$$\hat{\sigma}_b^2 = [\hat{b}'\hat{b} + \hat{\sigma}_e^2 \text{tr}(C^{22})]/q$$

$$\hat{\sigma}_g^2 = [\hat{g}'G^{-1}\hat{g} + \hat{\sigma}_e^2 \text{tr}(C^{33})]/s$$

$$\hat{\sigma}_i^2 = [\hat{i}'\hat{i} + \hat{\sigma}_e^2 \text{tr}(C^{44})]/t$$

Onde:

C^{22} , C^{33} e C^{44} advêm da inversa da matriz dos coeficientes das equações do modelo misto, isto é:

$$C^{-1} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ C_{41} & C_{42} & C_{43} & C_{44} \end{bmatrix}^{-1} = \begin{bmatrix} C^{11} & C^{12} & C^{13} & C^{14} \\ C^{21} & C^{22} & C^{23} & C^{24} \\ C^{31} & C^{32} & C^{33} & C^{34} \\ C^{41} & C^{42} & C^{43} & C^{44} \end{bmatrix}$$

tr: operador traço matricial;

$r(x)$: posto da matriz X;

N, q, s: número total de dados, número de blocos, número de genótipos e número de combinações genótipo x ambiente, respectivamente.

Por meio desse modelo foram obtidos os preditores BLUP empíricos (eBLUP ou REML/BLUP) dos valores genotípicos livres da interação, dados por $\hat{\mu} + \hat{g}_i$, em que $\hat{\mu}$ é a média de todos os locais e $\hat{g}_i$ é o efeito genotípico livre da interação genótipo x ambiente. Para cada local j, os valores genotípicos (VG) são preditos por $\hat{\mu}_j + \hat{g}_i + (\hat{g}e)_{ij}$, em que $\hat{\mu}_j$ é a média do local j, $\hat{g}_i$ é o efeito genotípico do genótipo i, no local j, e $(\hat{g}e)_{ij}$ é o efeito da interação genótipo x ambiente relativo ao genótipo i.

A predição dos valores genotípicos capitalizando a interação média (gem) nos diferentes locais é dada por $\hat{\mu}_j + \hat{g}_i + \hat{g}e_m$, sendo calculada por: $\hat{\mu} + \left[\frac{\hat{\sigma}_g^2 + \frac{\hat{\sigma}_i^2}{n}}{\hat{\sigma}_g^2} \right] \hat{g}_i$, em que, $\hat{\mu}$ é a média geral de todos os locais, n é o número de locais e $\hat{g}_i$ é o efeito genotípico do genótipo i.

Quanto menor for o desvio padrão do comportamento genotípico através dos locais, maior será a média harmônica de seus valores genotípicos através dos locais. Assim a seleção pelos maiores valores da média harmônica do valor genotípico (MHVG) implica simultaneamente na seleção para produtividade e estabilidade. Adicionalmente, o uso da transformação $1/y$ considera também a instabilidade dentro dos locais.

$$MHVG_i = \frac{n_i}{\sum_{j=1}^{n_i} \frac{1}{VG_{ij}}}$$

Em que:

n_i : número de ambientes em que o genótipo i foi avaliado;

VG_{ij} : valor genotípico predito do genótipo i no ambiente j

Em termos de adaptabilidade, uma medida simples e eficiente no contexto dos modelos mistos refere-se à performance relativa dos valores genotípicos (PRVG) através dos ambientes. Neste caso, os valores genotípicos preditos (ou os dados originais) são expressos como proporção da média geral de cada local e, posteriormente, obtém-se o valor médio desta proporção através dos locais. Genericamente, a performance relativa tem sido usada há um longo tempo (WRIGHT et al. 1966, apud RESENDE, 2007) em termos de dados fenotípicos, e constitui a base do método de Annicchiarico (1992 apud RESENDE, 2007).

$$PRVG_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{VG_{ij}}{\mu_j}$$

Em que:

n_i : Número de ambientes nos quais o genótipo i foi observado;

μ_j : média do ambiente j ;

VG_{ij} : valor genotípico predito do genótipo i no ambiente j ;

VG_{ij} : $\mu_j + g_i + ge_{ij}$, em que g_i é o efeito genotípico do genótipo i e o ge_{ij} é o efeito da interação genótipo x ambiente do genótipo i no ambiente j , ambos efeitos estimados via REML/BLUP.

A seleção simultânea por produtividade, estabilidade e adaptabilidade, no contexto de modelos mistos, pode ser realizada pelo método da média harmônica da performance relativa do valor genotípico (MHPRVG) predito, proposto por Resende (2004):

$$MHPRVG_i = \frac{n_i}{\sum_{j=1}^{n_i} \frac{1}{PRVG_{ij}}}$$

A estimação dos componentes de variância e efeitos de predição não viesada foram estimados via PROC MIXED, pelo método REML/BLUP utilizando-se o software SAS (SAS INSTITUTE INC., 2010).

2.3. Norma de reação com covariáveis desconhecidas (SU et al., 2006)

Encontra-se no CAPÍTULO 3 - Reaction norm with unknown covariates and Eberhart-Russell models to sugarcane's multi-environment trial.

3. REFERÊNCIAS BIBLIOGRÁFICAS

- ANNICCHIARICO, P. Cultivar adaptation and recommendation from alfafa trials in Northern Italy. **Journal of Genetics & Breeding**, Berlin, v. 46, n. 1, p. 269-278, 1992.
- CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2 ed., 2001. 390p.
- EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, n. 1, p. 36-40, 1966.
- FINLAY, K.W.; WILKINSON, G.N. The analysis of adaptation in a plant breeding 24 programme. **Aust. J. Agri. Res.**, 14, 742-754, 1963.
- RESENDE, M. D. V. **Métodos Estatísticos Ótimos na Análise de Experimentos de Campo**. Colombo: Embrapa Florestas, Ed. 1, 2004, 57 p. (Documentos, 100).
- RESENDE, M.D.V. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas. 2007. 561p.
- SAS Institute Inc. 2010. SAS/STAT® 9.22 User's Guide. Cary, NC: SAS Institute Inc.
- SU, G.; MADSEN, P.; LUND, M. S.; D, SORENSEN; KORSGAARD, I.R.; JENSEN, J. Bayesian analysis of the linear reaction norm model with unknown covariates. **Journal of Animal Science**, v. 84, p. 1651–1657, 2006.
- WRIGHT, J.W.; PAULEY, S.S.; POLK, R.B.; JOKELA, J.J. Performance of Scotch pine varieties in North Central Region, **Silvae Genetica**, v. 15, 0. 101-110, 1966.

CAPÍTULO 3 - Reaction norm with unknown covariates and Eberhart-Russell models to sugarcane's multi-environment trial

****Capítulo 3 escrito de acordo com as normas da Revista Scientia Agricola**

ABSTRACT

Multi-environment trials are commonly conducted in plant breeding programs to evaluate potential genotypes mainly to yield. This study included 16 sugarcane genotypes at 14 locations evaluated in four years, that is, four harvests (sugarcane plant, first, second and third ratoon). The aim was to apply and investigate the adaptability and stability for sugarcane genotypes for Ton of Pol per Hectare (TPH) in a multi-environment trial by the classical method of linear regression and the reaction norm with unknown covariates. The percentage of ranking coincidence from both approach was 31, 44, 50 and 56% at the first (sugarcane plant), second, third and fourth harvest respectively. Both methodologies assist the allocation of the genotypes in the environments available for planting and to establish breeding objectives. In sugarcane, the TPH attribute heritability is dependent of environmental gradient and increase along the harvests.

1. INTRODUCTION

Sugarcane is the most applied crop in the sugar yielding as much as internal consumption as exportation product in Brazil, besides its application in yielding alcohol and biomass production. The crop area has been increased expressively leading breeders to search more yielding genotypes to the different planted areas.

Multi-environment trials are commonly conducted in plant breeding programs to evaluate potential genotypes mainly to yield. In quantitative genetics the presence of genotype by environment interaction ($G \times E$) implies that different genotypes will respond differentially to environmental changes, possibly leading to re-rankings of performance in different environments. This presumes that the performance of plants and their progenies can change according to the environmental conditions in which they are harvested and selected and this is investigated by introducing random $G \times E$ effects in a linear model, using reaction norm models (Falconer and Mackay, 1996; Lynch and Walsh, 1998). According to Weikai and Hunt (2001), the understanding of causes of $G \times E$ can be used to

establish breeding objectives, identify ideal test conditions, and formulate recommendations for areas of optimal variety adaptation.

Reaction norm was defined by Schmalhausen (1949) as the set of phenotypes that can be produced by an individual genotype exposed to different environmental conditions. The introduction of random regression models in longitudinal data studies (Kirkpatrick and Heckman, 1989; Kirkpatrick et al., 1990; Meyer, 1998) opened new perspectives in animal genetic evaluations, including the G×E assessment using an reaction norm approach (De Jong and Bijma, 2002; Schaeffer, 2004).

The reaction norm model is an interesting alternative to describe G×E partly because it can accommodate a large number of environmental levels with few parameters. It assumes that the response variable is linearly related to the covariate representing the environmental variable. The slope associated with a given individual is a measure of its environmental sensitivity, and the amount of variation in the population slope indicates the degree of G×E that exists for the trait under consideration. Linear reaction norm have usually been used to detect G×E because they are simple to interpret (Calus and Veerkamp, 2003; Kolmodin et al., 2002; Oseni et al., 2004).

In the reaction norm model the value of the covariate is typically unknown and an approximation used in the literature consists of replacing the covariate by the average phenotypic performance of the individuals in a given environment [e.g., herd-year (Fikse et al., 2003; Hayes et al., 2003; Kolmodin et al., 2002; Shariati et al 2007)]. Other suggestions are to estimate environmental effects in a standard additive genetic model and to use these as proxies for the unknown covariates in the reaction norm model (Calus et al., 2002; Oseni et al., 2004), or using the reaction norm model, to replace the covariates by estimates obtained in an iterative manner (Calus et al., 2004). These methods have a number of shortcomings, especially when the trait of interest is under selection (Calus et al., 2004; Su et al., 2006).

Su et al. (2006) developed a Bayesian reaction norm model for inferring genetic parameters and environmental covariates simultaneously. The advantage of this approach over the approximate methods was shown in their simulation study.

Eberhart and Russell (1966) proposed a joint linear regression model where the phenotypic mean of each genotype is regressed as a function of an Environmental Index (I_j) which is expressed by difference between phenotypic mean of all genotypes in each environment and the general phenotypic mean of all genotypes in all environments. Thus,

the phenotypic product of each genotype is expressed by a straight line, $y = \alpha + bI_j$, which intercept α express the phenotypic product at $I_j=0$ and the slope b express the equivalent phenotypic reaction norm effect. In this paper the least square estimation of α and b is simultaneously applied to all genotypes and I_j . The observed I_j were compared with the ones produced by DMU package (Madsen and Jensen, 2004) estimated via reaction norm model.

The aim of this study was to apply and investigate the adaptability and stability of sugarcane genotypes for Ton of Pol per Hectare attribute in a multi-environment trial by the classical method of joint linear regression and the reaction norm with unknown covariates.

2. MATERIAL AND METHODS

2.1. Sugar cane data set

The data set were collected from trials of Procana IAC, installed in 2007, concerning four harvests from 2008 to 2011 (cane plant, first, second and third ratton). One evaluated 16 genotypes in a total of 14 mills (environments), where one had 13 environments to the first harvest (sugarcane plant), 12 environments to the second harvest, 10 to the third harvest and seven to the fourth harvest. The environments find themselves describe in the CAIANA software under dominion of IAC (Ribeirão Preto, São Paulo State, Brazil). The attributes measured were the ton of stalks per hectare (TSH) and the apparent percentage of sucrose (POL%) transformed into only one attribute, Ton of Pol per Hectare (TPH), by the formula: $TPH = (TSH * POL\%)/100$. The TPH was the attribute submitted to analysis. The design applied to each environment was complete random blocks with three replications. The plot was composed of five rows with ten meters and the space between rows was 1.5 m. The TPH was measured in nine testing genotypes: IACSP952289, IACSP962008, IACSP962073, IACSP963048, IACSP963076, IACSP972084, IACSP985024, IACSP985046 and IACSP986202; and in seven standard varieties: CTC2, CTC4, IACSP933046, RB72454, RB867515, RB92579 and SP813250. The standard varieties choice was a function of the adaptability and stability displayed by the season and environment's region where they were evaluated previously as well as the sugarcane crop representativeness area.

2.2. Linear reaction norm model with unknown covariates

The model used in this work is the same as described by Su et al. (2006). They proposed a Bayesian approach for jointly inferring upon all unknowns, h_j , β , g_i , and b_i , in the following One-step linear reaction norm model:

$$y_{ijk} = \mu + g_i + b_i h_j + h'_j + \beta_k(h_j) + e_{ijk}$$

This model can be rewritten for all n observation in matrix notation as follows:

$$y = \mathbf{X}\beta + \mathbf{W}h + \mathbf{Z}_\alpha g + \mathbf{Z}_i b + e$$

$y = \{y_{ijk}\}$ is the $n \times 1$ vector of observations (TPH attribute), β is the vector of order p (block nested within environment), $h = \{h_j\}_{j=1}^{n_h}$ the vector of environmental effects (alike to Environmental Index (I_j) from Eberhart-Russell), $g = \{g_i\}_{i=1}^q$ is the of random genotypic intercepts, $b = \{b_i\}_{i=1}^q$ the vector of slopes for genotype reaction norm, and e is the $n \times 1$ vector of residuals. The $\mathbf{X}$, $\mathbf{W}$ and $\mathbf{Z}_\alpha$ are known incidence matrices, whereas $\mathbf{Z}_i$ is inferred incidence matrix (reaction norm of the genotypic variance of each genotype). In other words, properly recognizing that h is unknown, the row of $\mathbf{Z}_i$ that corresponds to the same row address or element y_{ijk} in y is all zeroes except for h_j in column j (Su et al., 2006).

The conditional distribution of y is assumed to be normal having the form

$$y|\beta, h, g, b, \mathbf{R} \sim N(\mathbf{X}\beta + \mathbf{W}h + \mathbf{Z}_\alpha g + \mathbf{Z}_i b, \mathbf{R})$$

Where $\mathbf{R}$ is the matrix (order n) of random residual covariances. Without loss of generality, it is assumed that residuals are homocedastic and independent of each other, so that $\mathbf{R} = \mathbf{I}\sigma_e^2$, where $\mathbf{I}$ is the identity matrix ($n \times n$) and σ_e^2 is the residual variance.

2.3. Prior Specifications

The Gibbs sampler is a Monte Carlo method for obtaining samples from joint or marginal posterior distributions of all parameters in the model, by repeated sampling from fully conditional posterior distributions. The Gibbs sampling algorithm was implemented using the DMU package (Madsen and Jensen, 2004) cited in Su et al. (2006).

The prior distribution of β is assumed to be improper uniform. The random vectors g and b assumed to have a normal, mutually independent prior distributions.

The prior distributions of the variance of h_i (σ_h^2) and the residual variance (σ_e^2) are

assumed to be scaled inverse χ^2 distributions.

$$\begin{aligned}\beta &\sim p(\beta) \\ h \mid \sigma_h^2 &\sim N(0, \mathbf{I}\sigma_h^2) \\ \begin{bmatrix} g \\ b \end{bmatrix} &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_g^2 & \sigma_{gb} \\ \sigma_{gb} & \sigma_b^2 \end{bmatrix} \otimes \mathbf{A}\right)\end{aligned}$$

Where $p(\beta)$ is proportional to 1, σ_h^2 is the environmental effect variance, σ_g^2 and σ_b^2 are, respectively, the genotypic variances due to the reaction norm intercept and slope, and σ_{gb} is the genotypic covariance between reaction norm intercept and slope. The relationship matrix $\mathbf{A}$ was assumed as an identity matrix in this paper.

2.4. Bayesian Inference

The Gibbs sampling (Monte Carlo method) was used for obtaining samples from joint or marginal posterior distributions of all parameters in the model, by repeated sampling from fully conditional posterior distributions, all the fully conditional distributions are of standard form and are given in Su et al. (2006). The Gibbs sampling algorithm was implemented using the DMU package (Madsen and Jensen, 2004).

The chain length was 110,000 iterations, and the initial burn-in period was 10,000 iterations, generating posterior samples every 10 cycles. The convergence of the chains was evaluated by trace plot analysis of the values of the samples for all (co)variance components vs. the chain interaction using the package *boa* (Smith, 2007) from the R software (R Core Team, 2012). (Apêndice).

2.5. Variance components and Genetic parameters

According to Cardoso and Tempelman (2012), using the environmental gradient information, the genotypic value (GV) of a genotype i was estimated for each environmental level using the following formula:

$$GV_i \mid h_j = g_i + b_i h_j$$

Such as the genotypic variance for a given environment j with effect h_j , which we describe as $\sigma_G^2 \mid h_j$ can be obtained as follows:

$$\sigma_G^2 \mid h_j = \text{var}(g_i + b_i h_j) = \sigma_g^2 + h_j^2 \sigma_b^2 + 2h_j \sigma_{gb}$$

Where σ_g^2 is the genotypic variance, σ_b^2 is the genotypic slope variance and σ_{gb} is the covariance of the genotypic level and slope. Therefore, the specific heritability to environment j is determined as $h_G^2 | h_j = \frac{\sigma_G^2 | h_j}{\sigma_G^2 | h_j + \sigma_e^2 | h_j}$, where $\sigma_e^2 | h_j$ is the residual variance in environment j , which is constant, i.e., $\sigma_e^2 | h_j = \sigma_e^2 \forall i$ for the homoskedastic errors.

2.6. Eberhart-Russell Regression Analysis

The statistical model used was:

$$y_{ijk} = \mu + g_i + b_i I_j + D_i + \beta_k(I_j) + e_{ijk}$$

where $\mu + g_i$ is the phenotypic mean of the genotype i in environment $I_j=0$, b_i is the slope ("reaction norm") in function of the environment, D_i the linear regression deviations, $\beta_k(I_j)$ the block effect within environment j . That is, one proceed a simultaneous estimation on all environments with the minimization of e_{ijk} . To adaptability on all environments one might expect D_i not meaningful ($p > 0.05$). Otherwise the genotype not unchangeably adapt in all environments. The stability is expected with $b_i=1$; if $b_i < 1$ denotes not responsively to the environment (rustic genotype) e $b_i > 1$ denotes greater response. In general, genotypes with $b_i > 1$ are indicate to better environments (responsive genotypes). That is, Eberhart and Russell (1966) emphasized the need of considering both linear (b_i) and not meaningful (D_i) components of genotype environment interactions in judging the stability of a genotype. If D_i is meaningful ($p < 0.05$), b_i becomes not estimable to practical purposes (solution nonsense).

To simultaneous estimation of g_i and b_i and the meaningful of D_i it was built a SAS program via PROC GLM (SAS Institute, 2010).

2.7. Reaction norm and Breeding

Practically, the response (Y) in function of an environmental index (X) often has been observed as linear. One might consider from the mathematic point of view, supposing Y and X are normal variables (that is, binormality) it's demonstrated the expectation of Y given X as a straight line, $E(Y | X) = g + bX$, the parameters g and b depends on the means, variances and (co)variances of Y and X (Hogg and Craig, 1965, p. 103-104).

Thus, one of the enough conditions for the straight line be representative is the responses (Y) and the environmental indexes (X) be binormals. This happens with genotypes that adapt to the environments, responding worst or better according the environment gets better or worst. Genotypes that show this property are the ones with adaptability, or with plasticity to several environments. When the genotype doesn't show such property, the linear regression deviances become meaningful and the genotypes only must be selected to specific environments.

3. RESULTS AND DISCUSSION

The Eberhart-Russell joint linear regression model (E-R) has been widely used in the past few decades mainly because variability in performance of any genotype could be subdivided into predictable (regression) and unpredictable ($MS-D_i$) components. Hypothetically, the E-R method considers both yield (regression) and stability ($MS-D_i$), with regression being predicted and to some extent controlled by selecting specific genotypes for specific locations.

The E-R joint linear regression analysis results are in Table 1 and 2. According to E-R method, the genotypes with high regression coefficients are considered as responsive yielding genotypes and those with low regression coefficients as nonresponsive yielding. At the same time, the genotypes with high $MS-D_i$ estimates are considered as highly unstable and those with low $MS-D_i$ estimates as highly stable. To the first harvest (Table 1), the genotypes CTC2 and IACSP963048 are considered responsive due to their significant Regression Coefficient (RC) and stable due to their non-meaningful $MS-D_i$. Still in the first harvest, the genotypes CTC4, IACSP933046, IACSP962008, IACSP962073, IACSP963076, IACSP972084, IACSP985046, RB72454 and RB867515 find themselves as stables due to the not meaningful $MS-D_i$. To the others genotypes that the $MS-D_i$ were meaningful such as IACSP952289, IACSP985024, IACSP986202, RB92579 and SP813250 one might say these ones have preferential adaptability, that is, their response would be responsive and acceptable only to specific environments. Thus, there is no meaning in estimate RC once it would be nonsense. According to Alwala et al. (2010), increased GxE is usually associated with decreased correlation between genotypic and phenotypic values which in turn reduces gain from selection. These genotypes have a greater and specific interaction with only few environments and, in general, are not

desirable to breeders and agronomists. Moreover, if the RC are not significantly different from each other, the slopes may be ignored and only the yield mean and $MS-D_i$ could be used as parameter to classify genotypes by stability and yielding. Therefore, at the first harvest, the E-R results show us that CTC2, IACSP962073, IACSP963048, IACSP963076 and RB867515 were the top five recommendable genotypes to be planted in this environment set.

Following the second harvest (Table 1), one might observe via E-R analysis a different genotype disposition, where the responsive and stability genotypes were IACSP962073, RB867515 and SP813250 due to their significant RC and nonlinear $MS-D_i$. From a non-meaningful $MS-D_i$ one might have the CTC4, IACSP933046, IACSP952289, IACSP962008 and IACSP985024 as stables. The genotypes which showed preferential adaptability were CTC2, IACSP963048, IACSP963076, IACSP972084, IACSP985046, IACSP986202, RB72454 and RB92579. Thus, the top five recommendable genotypes to be planted in this environment set were CTC4, IACSP962073, IACSP985024, RB867515 and SP813250.

Table 1. Mean yield, mean yield ranks and results from Eberhart-Russell joint regression analysis to Ton of Pol per Hectare attribute in a sugarcane's multi environment trial at the first and second harvests.

Genotype	First harvest				Second harvest			
	Rank	μ	RC	$MS-D_i$	Rank	μ	RC	$MS-D_i$
CTC2 ^{SV}	2	20.13	1.10**	7.55 ns	1	17.02	-	13.20**
CTC4 ^{SV}	13	17.35	-37.97 ns	12.66 ns	2	16.20	1.20 ns	2.85 ns
IACSP933046 ^{SV}	12	17.53	0.74 ns	3.13 ns	11	14.57	4.17 ns	2.45 ns
IACSP952289	15	15.68	-	8.86**	13	14.25	0.59 ns	4.01 ns
IACSP962008	10	17.81	2.58 ns	7.82 ns	10	14.85	0.92 ns	5.29 ns
IACSP962073	8	18.03	0.30 ns	8.28 ns	7	15.30	1.24*	1.11 ns
IACSP963048	6	19.08	3.68**	13.33 ns	5	15.72	-	15.31**
IACSP963076	4	19.64	-0.59 ns	17.88 ns	3	15.78	-	1.67**
IACSP972084	16	14.84	-0.37 ns	6.95 ns	16	12.16	-	5.37**
IACSP985024	7	18.33	-	11.74*	9	15.04	0.50 ns	8.93 ns
IACSP985046	14	16.91	0.68 ns	7.43 ns	12	14.41	-	6.97*
IACSP986202	11	17.59	-	11.51**	15	13.45	-	12.84**
RB72454 ^{SV}	9	17.93	1.56 ns	7.94 ns	14	14.03	-	10.86**
RB867515 ^{SV}	5	19.39	0.11 ns	9.67 ns	4	15.74	1.17*	4.44 ns
RB92579 ^{SV}	1	21.16	-	29.23**	6	15.63	-	11.76**
SP813250 ^{SV}	3	20.01	-	13.85*	8	15.29	1.71*	7.18 ns
General mean		18.21				14.96		

^{SV}: Standard variety; μ : phenotypic mean; RC: Regression Coefficient; $MS-D_i$: Mean Square of the linear regression deviation; *: $p < 0.05$; **: $p < 0.01$; ns: $p > 0.05$; - nonsense solution

To the third harvest (Table 2), one might have a different disposition even from the first and second harvests. Where the best genotypes to keep in the breeding program according to E-R analysis is the ones with RC meaningful and MS-D_i not meaningful and high yielding, due to their prediction in response and stability to keep such response. A very responsible genotype was IACSP963048, in other hand, one had IACSP952289. Both of these genotypes showed strong response to the environmental gradient changes which not means they were the better ones. As stables but with not enough meaningful statistical prediction response are IACSP962008, IACSP962073, IACSP985024, IACSP985046, IACSP986202, RB72454, RB867515, RB92579 and SP813250. The ones with preferential adaptability/response, that is, the ones that had showed MS-D_i meaningful as one can observe is CTC2, CTC4, IACSP933046, IACSP963076 and IACSP972084. Hence, one might have the top five recommendable genotypes would be those ones based on their non-meaningful MS-D_i and yielding such as IACSP962008, IACSP962073, RB867515, RB92579 and SP813250 at the third harvest.

To the fourth harvest (Table 2), one might find some similar genotypes as in the first harvest (such as CTC2 and IACSP963048) that were responsive and stable as well as CTC4, IACSP933046, IACSP962073, IACSP985024 and RB92579 due to their RC meaningful and MS-D_i non-meaningful or tending zero. The genotypes that not showed RC meaningful and neither MS-D_i were the stables ones but with no ability to predict the response behavior through the RC (IACSP962008, IACSP985046, IACSP986202, RB867515 and SP813250). The genotypes that showed preferential adaptability likewise were IACSP952289, IACSP963076, IACSP972084 and RB72454 genotype. What might lead us to have some potential genotype to specific environments, once this specific adaptation genotypes might be chosen to cross to others adaptable and stable ones in order to get a higher frequency of favorable genes to yielding in a larger number of environments. Hence the recommendable genotypes to the fourth harvest in this environment set were CTC2, IACSP962008, IACSP962073, IACSP985024 and RB867515.

Table 2. Mean yield, mean yield ranks and results from Eberhart-Russell joint regression analysis to Ton of Pol per Hectare attribute in a sugarcane's multi environment trial at the third and fourth harvests.

Genotype	Third harvest				Fourth harvest			
	Rank	μ	RC	MS- D_i	Rank	μ	RC	MS- D_i
CTC2 ^{SV}	8	13.18	-	6.28**	1	13.46	1.46*	9.95ns
CTC4 ^{SV}	3	14.32	-	7.22*	8	10.92	1.44**	7.39ns
IACSP933046 ^{SV}	5	13.55	-	2.66**	7	11.13	1.21**	1.49ns
IACSP952289	12	11.89	-8.35*	3.65 ns	13	9.52	-	5.10*
IACSP962008	7	13.41	-4.77 ns	3.00 ns	6	11.21	0.93 ns	1.55ns
IACSP962073	4	13.72	-5.14 ns	3.10 ns	3	11.89	3.14**	4.62ns
IACSP963048	14	11.71	6.95*	3.16 ns	11	10.24	1.63**	1.13ns
IACSP963076	6	13.47	-	10.01**	2	12.50	-	9.21*
IACSP972084	16	11.31	-	7.06**	15	9.09	-	5.94**
IACSP985024	11	12.34	16.13 ns	7.41 ns	4	11.68	2.53*	2.32ns
IACSP985046	13	11.84	2.57 ns	2.97 ns	12	10.13	-0.48 ns	5.30ns
IACSP986202	15	11.36	-4.02 ns	5.64 ns	14	9.13	0.80 ns	0.78ns
RB72454 ^{SV}	10	12.63	-0.20 ns	1.72 ns	16	8.99	-	8.83**
RB867515 ^{SV}	1	15.29	-7.81 ns	5.16 ns	5	11.66	0.42 ns	1.48ns
RB92579 ^{SV}	2	14.91	0.91 ns	0.33 ns	9	10.72	1.19**	0.17ns
SP813250 ^{SV}	9	13.10	-2.86 ns	1.63 ns	10	10.71	-0.05 ns	1.80ns
General mean		13.00				10.81		

^{SV}: Standard variety; μ : phenotypic mean; RC: Regression Coefficient; MS- D_i : Mean Square of the linear regression deviation; *: $p < 0.05$; **: $p < 0.01$; ns: $p > 0.05$; - nonsense solution

The reaction norm linear regression (RN) analysis results are in Table 3 and 4. Following some different principle from E-R method, the RN method show us that genotypes with regression coefficient (b_i) meaningful and higher than zero are the responsible genotypes. The genotypes with b_i non-meaningful are the ones that are stables and the genotypes with b_i meaningful and lower than zero are the ones that respond less than the mean of all others according the increase of the environment gradient.

To the first harvest (Table 3), the genotypes with meaningful and higher than zero b_i are CTC2 and IACSP3048 (the same to E-R method, Table 1). Nevertheless the RN method identifies the IACSP952289 and IACSP972084 as genotypes that would yield better when find themselves in unfavorable environments due their meaningful and lower than zero b_i (- 0.126 and - 0.185, respectively). The others genotypes evaluated show themselves as stables. Thus, from the RN results one might have as top five best recommendable genotype the CTC2, IACSP963048, IACSP963076, R867515 and

SP813250.

To the second harvest (Table 3), the IACSP972084 and IACSP986202 genotypes show themselves as genotypes recommendable to be planted in unfavorable environments due their meaningful and lower than zero b_i (- 0.293 and - 0.196, respectively). The CTC2 genotype show itself very responsible. It's is indeed a very good genotype however must be used carefully because its very sensibility to the environmental changes, that is, sensible even to yield more as well as to suffer losses when planted in unfavorable environments. The others genotypes showed their b_i non-meaningful, fact that lead us to label them as stable genotypes. Therefore one might find as the top five recommendable genotypes the CTC2, CTC4, IACSP963048, IACSP963076 and RB92579.

Table 3. Mean genotypic predicted value (MGV), MGV ranking and regression coefficient (b_i) from Reaction norm to Ton of Pol per Hectare attribute in a sugarcane's multi environment trial at the first and second harvests.

Genotype	First harvest			Second harvest		
	Rank	MGV	b_i	Rank	MGV	b_i
CTC2 ^{SV}	1	20.62	0.156 *	1	17.27	0.299 *
CTC4 ^{SV}	7	18.93	0.087 ns	2	16.16	0.143 ns
IACSP933046 ^{SV}	13	17.46	-0.121 ns	9	14.99	-0.031 ns
IACSP952289	15	15.81	-0.126 *	14	14.08	-0.152 ns
IACSP962008	11	17.82	-0.007 ns	11	14.72	-0.047 ns
IACSP962073	10	17.90	-0.091 ns	8	15.22	0.043 ns
IACSP963048	6	19.16	0.153 *	5	15.61	0.087 ns
IACSP963076	3	19.44	0.020 ns	3	15.76	0.138 ns
IACSP972084	16	14.99	-0.185 *	16	12.37	-0.293 *
IACSP985024	8	18.23	-0.034 ns	10	14.91	-0.018 ns
IACSP985046	14	16.93	-0.085 ns	12	14.31	-0.097 ns
IACSP986202	12	17.66	0.007 ns	15	13.46	-0.196 *
RB72454 ^{SV}	9	18.10	0.103 ns	13	14.11	-0.071 ns
RB867515 ^{SV}	4	19.26	0.039 ns	6	15.45	0.011 ns
RB92579 ^{SV}	5	19.23	0.022 ns	4	15.63	0.104 ns
SP813250 ^{SV}	2	19.83	0.062 ns	7	15.30	0.080 ns
General Mean		18.21			14.96	

^{SV}: Standard variety; *: $p < 0,05$; ns: non-meaningful.

To the third harvest (Table 4), one might observe once more as in the previous harvests that the IACSP972084 is a genotype recommendable to unfavorable environments ($b_i = -0.164^*$). Considering the genotypic value ranking and the b_i , the top five recommendable genotypes were CTC2, CTC4, IACSP933046, RB867515 and

RB92579.

To the fourth harvest (Table 4) all b_i were non-meaningful. This means that every genotypes were considered stable. The GxE interaction can be measure by the amount find in the b_i . In this case the selection of genotypes would be based only considering the genotypic value ranking. Where the ranking of top five recommendable genotypes would be CTC2, CTC4, IACSP962073, IACSP963076, and IACSP985024.

Table 4. Mean genotypic predicted value (MGV), MGV ranking and regression coefficient (b_i) from Reaction norm to Ton of Pol per Hectare attribute in a sugarcane's multi environment trial at the third and fourth harvests.

Genotype	Third harvest			Fourth harvest		
	Rank	MGV	b_i	Rank	MGV	b_i
CTC2 ^{SV}	3	14.42	0.212 *	1	13.40	0.007 ns
CTC4 ^{SV}	2	14.80	0.151 *	3	12.03	0.036 ns
IACSP933046 ^{SV}	5	13.84	0.035 ns	8	11.24	-0.016 ns
IACSP952289	12	11.77	-0.099 ns	13	9.44	-0.029 ns
IACSP962008	8	13.22	-0.117 ns	9	11.01	-0.038 ns
IACSP962073	6	13.51	-0.076 ns	4	11.68	0.076 ns
IACSP963048	14	11.62	0.029 ns	11	10.10	-0.050 ns
IACSP963076	7	13.30	0.111 ns	2	12.26	0.101 ns
IACSP972084	16	11.21	-0.164 *	15	9.00	-0.135 ns
IACSP985024	10	12.79	-0.023 ns	5	11.47	0.031 ns
IACSP985046	13	11.74	-0.015 ns	12	10.00	-0.045 ns
IACSP986202	15	11.29	0.035 ns	14	9.11	0.043 ns
RB72454 ^{SV}	11	12.48	-0.092 ns	16	8.95	-0.001 ns
RB867515 ^{SV}	1	15.00	-0.073 ns	6	11.45	0.011 ns
RB92579 ^{SV}	4	14.03	-0.073 ns	7	11.29	0.070 ns
SP813250 ^{SV}	9	12.96	0.159 *	10	10.53	-0.061 ns
General mean		13.00			10.81	

^{SV}: Standard variety; *: $p < 0,05$; ns: non-meaningful.

The Figures 1, 2, 3 and 4 show up the graphic of the genotypic value (GV) for the sixth highest genotypic values in the harvests analyzed (Figure a) as well as the sixth highest phenotypic means and general mean of all genotypes in each environment (Figure b). The ranking is showed in the legend of the respective figure. The GV may be interpreted as expected relative advantages regarding to the mean of the genotype cluster evaluated. It's important to plot the results of a genotype selection in multi environment trials based only in the genotypic value or/and only in the phenotypic mean. Once the breeders and the method user's might be able to see the implications of selection.

In Figure 1a (first harvest) one verify that the top five genotypes show up positive

GV from the -6 to 10 environmental gradient value, one highlight the genotype CTC2 with highest GVs. The genotypes SP813250, IACSP963076, RB867515 and RB92579 show up positive GVs and a potential to be the most adaptable ones. The IACSP963048 genotype show a slope (b_i) alike CTC2 genotype nevertheless not the same yielding. The top sixth genotypes selected by their highest GV showed positive slopes (b_i) one might observe the tendency to get higher GV as soon as the environmental gradient gets higher as well.

The Figure 1b shows the phenotypic means of the top six genotypes to first harvest based on their the phenotypic mean. One can observe the ranking does not match and the behavior of the genotypes seems not be stable through the environments. Selecting genotypes basing only in the phenotypic mean and not considering their regression coefficient may take the selection to a misleading in the recommendation. However, in the first harvest case, both approach indicates the same cluster of genotypes to be selected. In addition, a very good exploration of the result and a variety management may increase even more the yield of the selected genotype cluster. The percentage of ranking coincidence from both approach was 31%.

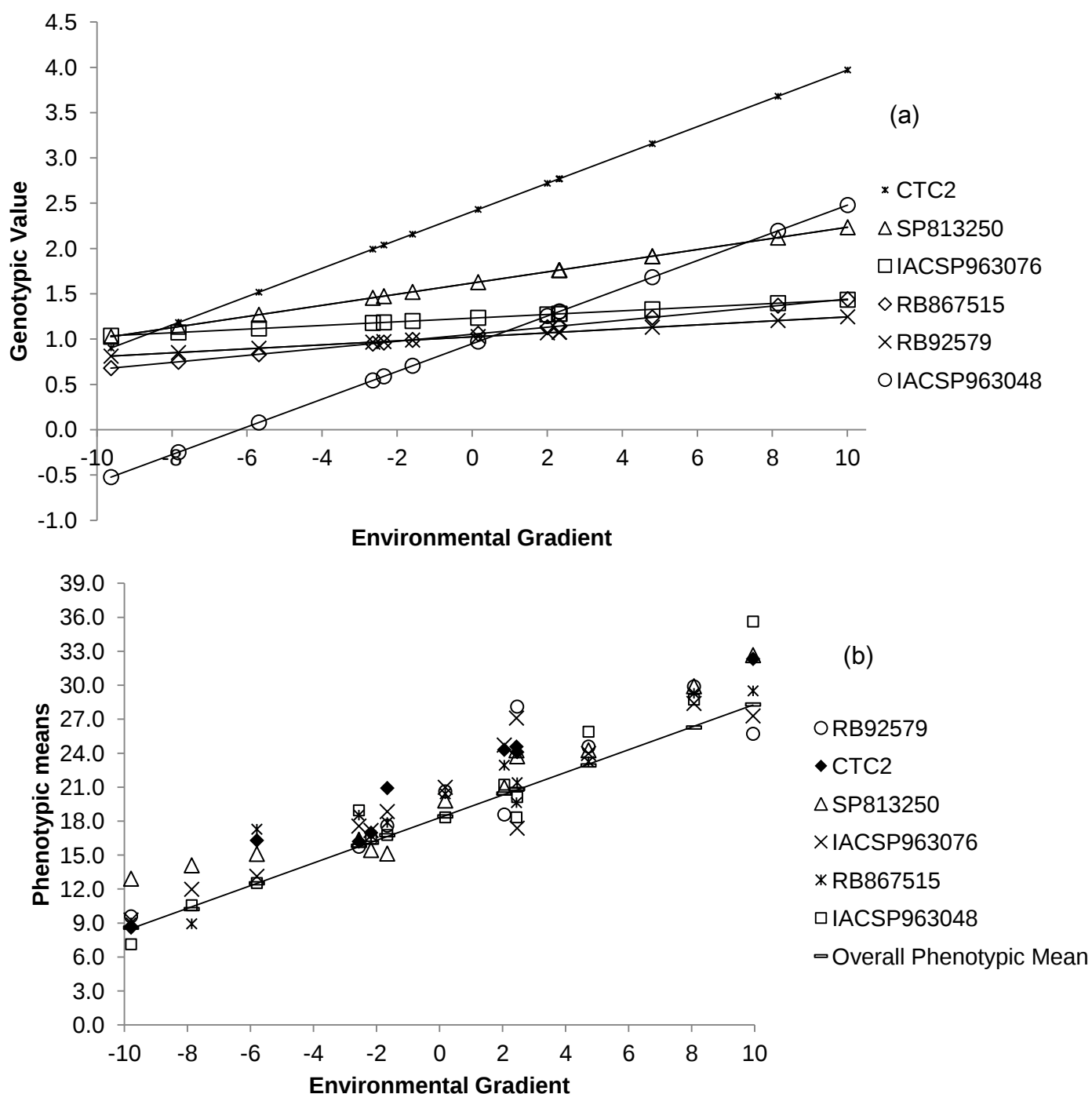


Figure 1. a) Reaction Norm along the environmental gradient for TPH obtained for the six highest genotypic values in the first harvest. b) Plot of the six highest phenotypic means and the overall mean along the environmental gradient to the first harvest.

In Figure 2a one might see the same behavior of the slopes of the six top highest GV selected as in Figure 1, that is, the slopes keep positive and the GV get higher as well as the environmental gradient value, except, of course, to the RB867515 genotype. The CTC2 and IACSP963076 genotypes remains in the rank. Only the three first genotypes

are the same in both approach. The Figure 2b shows how high is the top six genotypes yielding from the general phenotypic mean in each environment. The percentage of ranking coincidence from both approach was 44%.

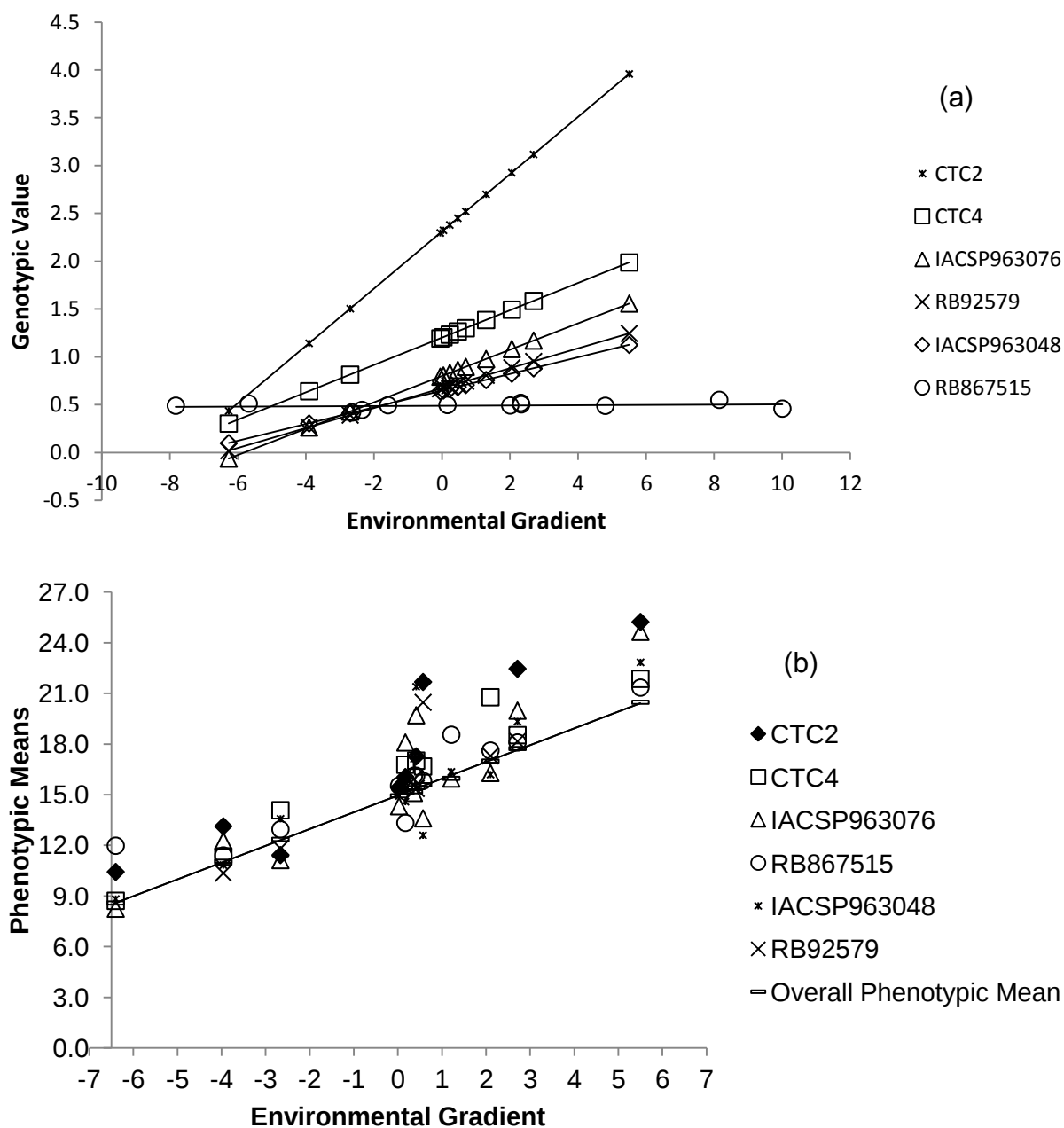


Figure 2. a) Reaction Norm along the environmental gradient for TPH obtained for the six highest genotypic values in the second harvest. b) Plot of the six highest phenotypic means and the overall mean along the environmental gradient to the second harvest.

The third harvest (Figure 3a) seems to have a very different behavior from the first and second harvests. One might observe the complex interaction based on the Figure 3a graphic where the top genotypes ranking in the environmental gradient -7.5 are not the

same as the genotypes ranking in the environmental gradient 6.5. Only CTC4, CTC2 and IACSP933046 slopes were positives. The RB867515, RB92579 and IACSP962073 slopes were negatives and may be selected to unfavorable environments as the most yielding genotypes based on TPH attribute analyzed. In phenotypic terms, according to Table 1, one might verify that RB867515 and RB92579 have been ranked as the first and second one by their phenotypic mean yield, nevertheless the Figure 3a shows that indeed they are good genotypes nevertheless to unfavorable environments. Meanwhile to CTC4 and CTC2 show better results as soon as the environmental gradient gets higher or more favorable. From this results the breeder could lead the genotype selection and the planting project in order to get the highest/best yield at each cluster of environments.

From the Figure 3b one might observe the behavior of the genotypes and the preferential adaptability of them, mainly to CTC4, IACSP933046 and IACSP963076. It highlights the issue in selection taking into account only the phenotypic mean. The percentage of ranking coincidence from both approach was 50% at the third harvest.

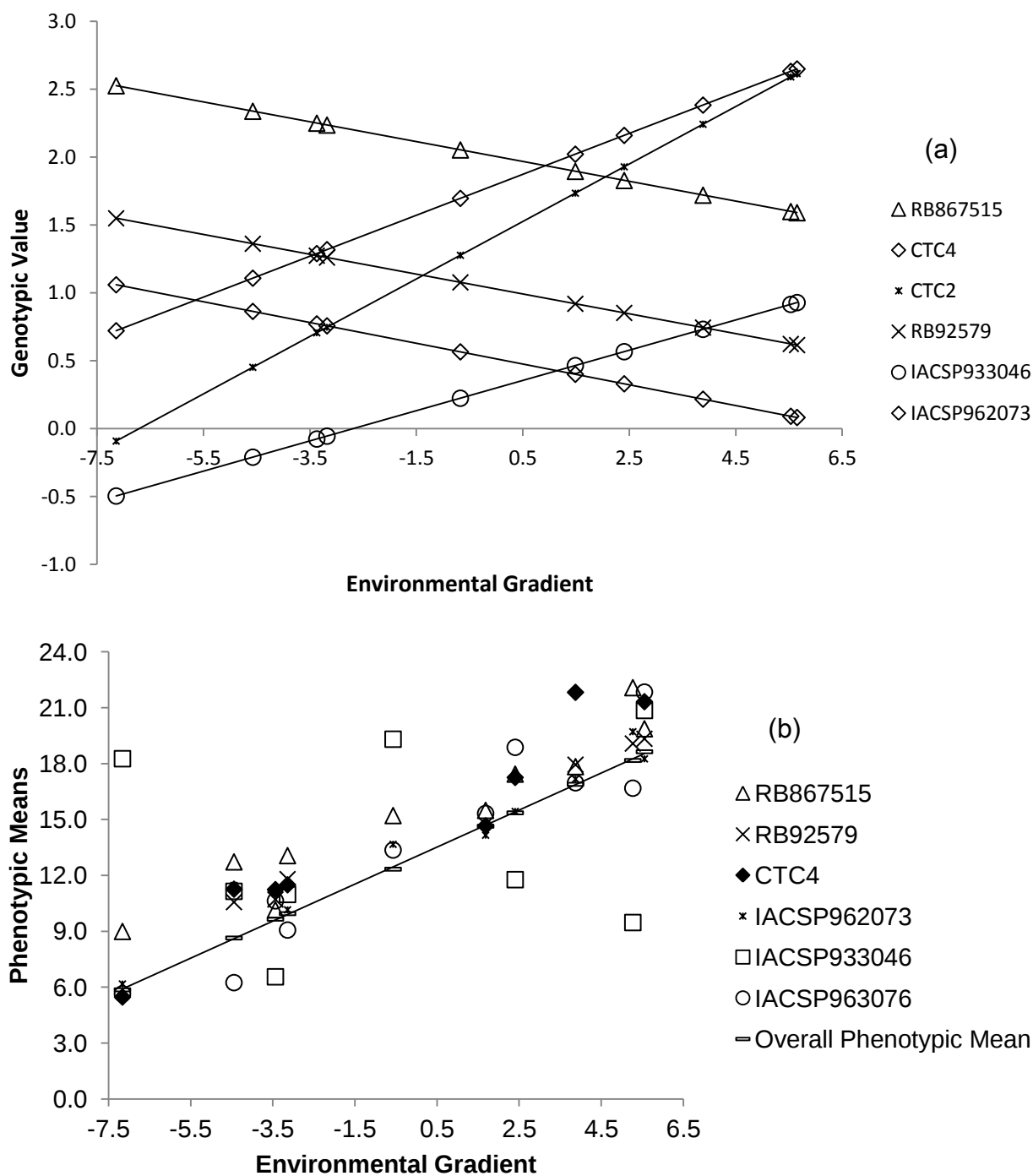


Figure 3. a) Reaction Norm along the environmental gradient for TPH obtained for the six highest genotypic values in the third harvest. b) Plot of the six highest phenotypic means and the overall mean along the environmental gradient to the third harvest.

To the fourth harvest (Figure 4a) all reaction norm slopes tending to be non-meaningful. It becomes the chosen of best genotypes only by the GV ranking. Differently from the Figure 4b which shows ideal genotypes from the E-R methodology point of view. The percentage of ranking coincidence from both approach to the fourth harvest was 56%.

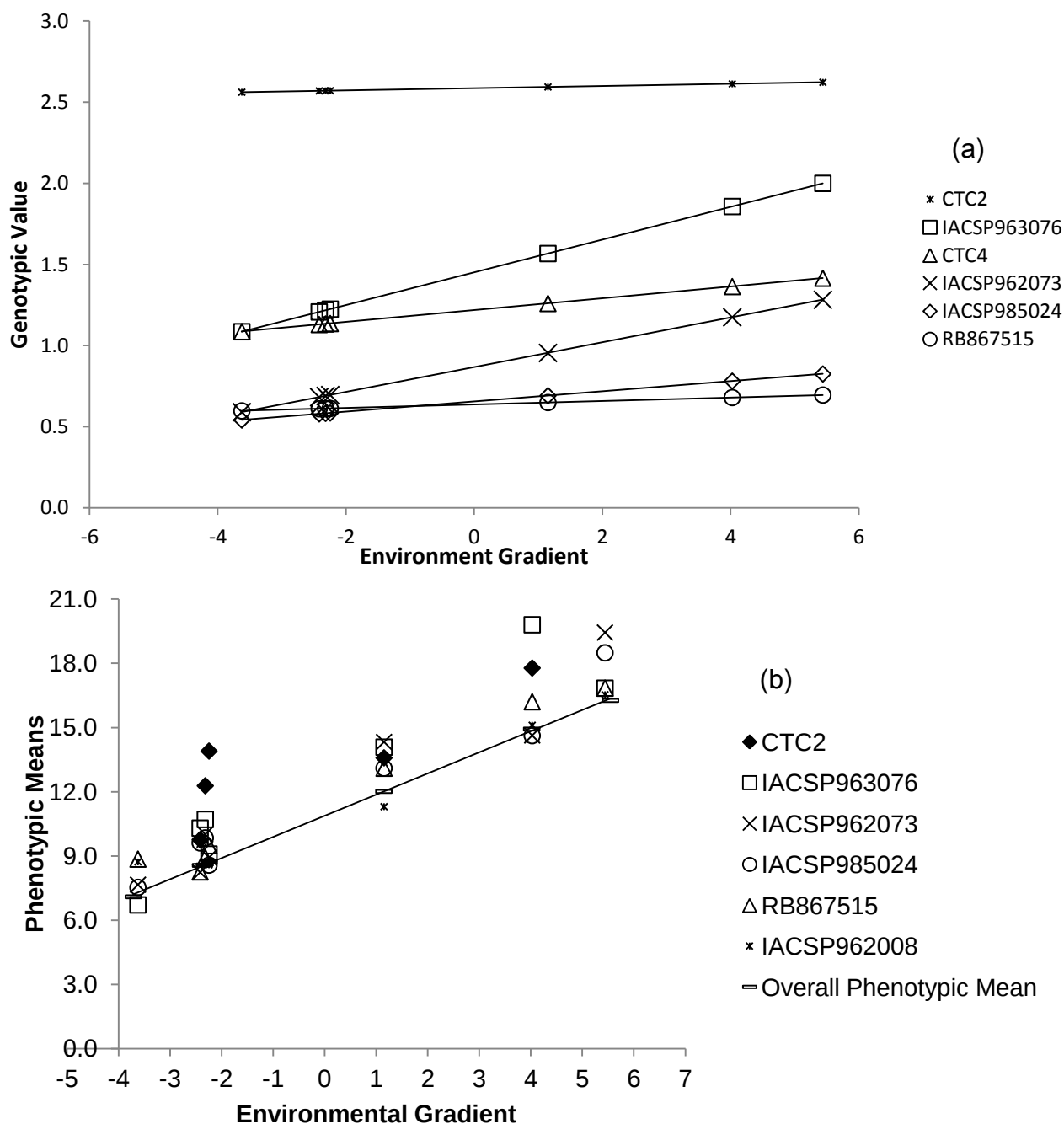


Figure 4. a) Reaction Norm along the environmental gradient for TPH obtained for the six highest genotypic values in the fourth harvest. b) Plot of the six highest phenotypic means and the overall mean along the environmental gradient to the fourth harvest.

From the Figure 5, the highest the environmental gradient the highest is the heritabilities to all harvests, that is, the genotypic variance and, consequently, the heritability increased along the environmental gradient. One can observe the third and fourth harvest, at greater levels of the environmental gradient, are the best ones to effectuate the selection for genetic varieties due to the highest heritabilities achieved and

thereby have better genetic gains along the breeding program.

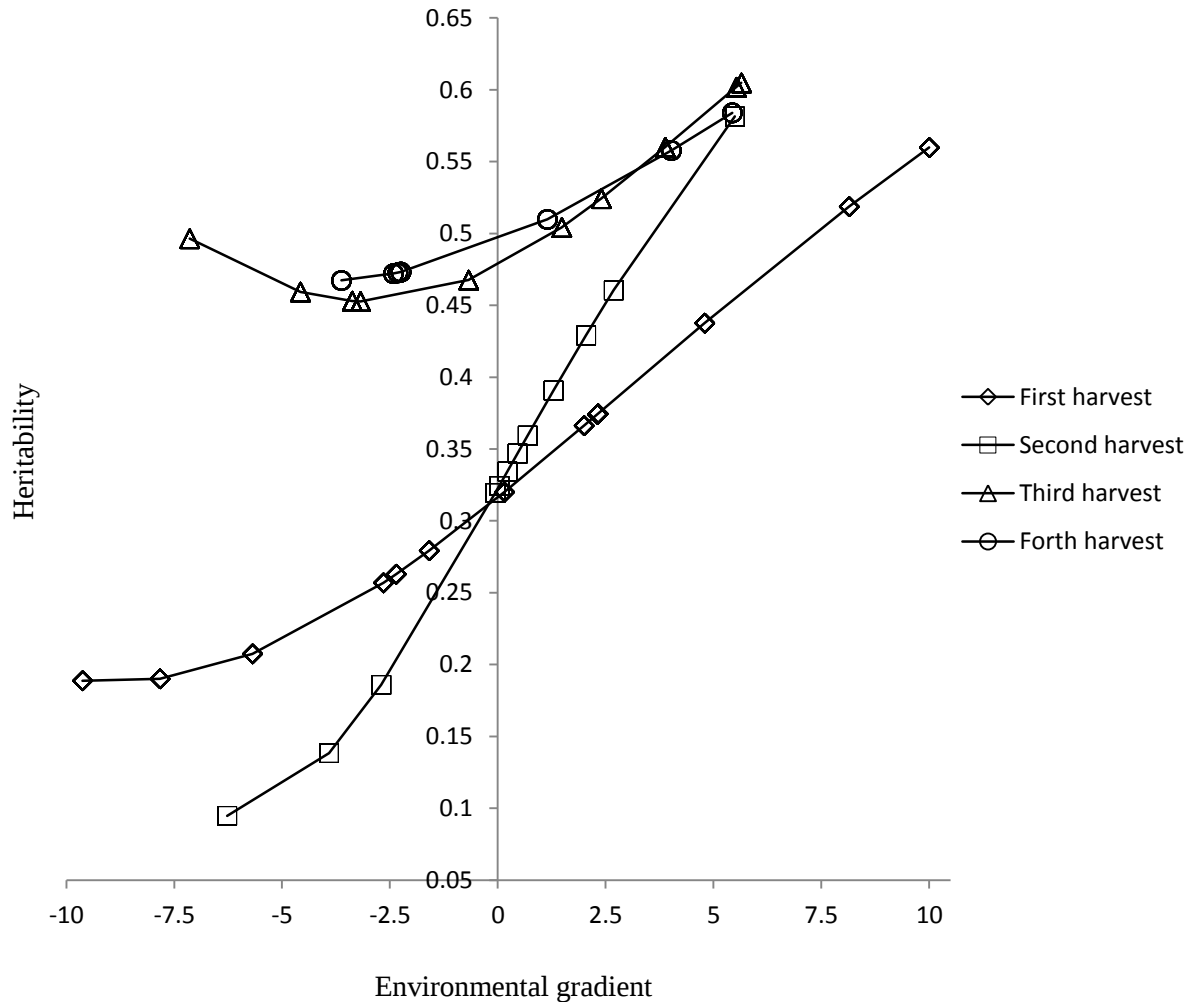


Figure 5. Heritabilities estimated for Ton Pol per Hectare attribute along the environmental gradient to four harvests in the sugarcane crop.

4. CONCLUSIONS

- The percentage of ranking coincidence from both approach was 31, 44, 50 and 56% at the first, second, third and fourth harvest respectively.
- Both methodologies assist the allocation of the genotypes in the environments available for planting and to establish breeding objectives.
- In sugarcane, the TPH attribute heritability is dependent of environmental gradient and increase along the harvests.

5. ACKNOWLEDGMENTS

The authors would like thank for both scholarship from the CAPES Foundation, Proc. no. BEX 5282/11-0 and the National Council of Technological and Scientific Development (CNPq) and the Professor Per Madsen from University of Aarhus, Tjele, Denmark.

6. REFERENCES

- Alwala, S.; Kwolek, T.; Mcpherson, M.; Pellow, J; Meyer, D. 2010. A comprehensive comparison between Eberhart and Russell joint regression and GGE biplot analyses to identify stable and high yielding maize hybrids. *Field Crops Research*, 119:225-230.
- Calus, M. P. L.; Bijma, P; Veerkamp, R. F. 2004. Effects of data structure on the estimation of covariance functions to describe genotype by environment interaction in a REACTION NORM model. *Genetics Selection Evolution* 36: 489–507.
- Calus, M. P. L.; Groen, A. F.; De Jong, G. 2002. Genotype $\times$ environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *Journal of Dairy Science* 85: 3115–3123.
- Calus, M. P. L.; Veerkamp, R. F. 2003. Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. *Journal of Dairy Science* 86: 3756–3764.
- Cardoso, F.F.; Tempelman, R.J. 2012. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction, *Journal of Animal Science* 90: 2130-2141.
- De Jong, G.; P. Bijma. 2002. Selection and phenotypic plasticity in evolutionary biology and animal breeding. *Livestock Production Science* 78:195–214.
- Eberhart, S. A.; Russell, W. A. 1966. Stability parameters for comparing varieties. *Crop Science*, Madison, 6:36-40.
- Falconer, D.S.; Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*. 4ed. Longman Group Ltd, Harlow, UK.
- Fikse, W. F.; Rekaya, R.; Weigel, K. A. 2003. Genotype $\times$ environment interaction for milk

- production in Guernsey cattle. *Journal of Dairy Science* 86:1821–1827.
- Hayes, B. J.; Carrick, M.; Bowman, P.; Goddard, M. E. 2003. Genotype $\times$ environment interaction for milk production of daughters of Australian dairy sires from test-day records. *Journal of Dairy Science* 86:3736–3744.
- Hogg, R.V.; Craig, A.T. 1965. *Introduction to mathematical statistics*, 2^a ed., New York, Macmillan Company. 388 p.
- Kirkpatrick, M.; Heckman, N. 1989. A quantitative genetic model for growth, shape and other infinite-dimensional characters. *Journal of Mathematical Biology* 27:429–450.
- Kirkpatrick, M.; Lofsvold, D.; Bulmer, M. 1990. Analysis of inheritance, selection and evolution of growth trajectories. *Genetics* 124:979–993.
- Kolmodin, R.; Strandberg, E.; Madsen, P.; Jensen, J.; Jorjani, H. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agriculturae, Scandinavica, Section A: Animal Science*. 52:11–24.
- Lynch, M.; Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates Inc., Sunderland, Massachusetts, USA.
- Madsen, P.; Jensen, J. 2004. *A User's Guide to DMU*, version 6, release 4.5. Danish Inst. Agric. Sci., Tjele, Denmark.
- Meyer, K. 1998. Estimating covariance functions for longitudinal data using a random regression model. *Genetics Selection Evolution* 30:221–240.
- Oseni, S.; Misztal, I.; Tsuruta, S.; Rekaya, R. 2004. Genetic components of days open under heat stress. *Journal of Dairy Science* 87:3022–3028.
- R Core Team (2012). *R. A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- SAS Institute Inc. 2010. *SAS/STAT® 9.22 User's Guide*. Cary, NC: SAS Institute Inc.
- Schaeffer, L. R. 2004. Application of random regression models in animal breeding. *Livestock Production Science* 86:35–45.
- Schmalhausen, I. I. 1949. *Factors of Evolution: The Theory of Stabilizing Selection*. The Blakiston Company, Philadelphia, PA.
- Shariati, M.M.; Su, G.; Madsen, P.; Sorensen, D. 2007. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. *Journal of Dairy Science*, 90: 5759-5766.
- Smith, B. J. 2007 *boa*: An R Package for MCMC Output Convergence Assessment and

Posterior Inference, *Journal of Statistical Software*, 21(11): 1-37

Su, G.; Madsen, P.; Lund, M. S.; Sorensen, D.; Korsgaard, I. R.; Jensen, J. 2006.

Bayesian analysis of the reaction norm model with unknown covariates. *Journal of Animal Science* 84:1651–1657.

Weikai, Y.; Hunt, L. A. 2001. Interpretation of genotype x environment interaction for winter wheat yield in Ontario. *Crop Science* 41: 19-25.

CAPÍTULO 4 - Média harmônica da performance relativa do valor genotípico em ensaios de múltiplos ambientes na cultura da cana-de-açúcar

RESUMO - A área de cultivo da cana-de-açúcar vem crescendo expressivamente, fato que leva melhoristas a buscar genótipos mais produtivos às diferentes áreas de plantio. Experimentos em multi ambientes são comumente conduzidos em programas de melhoramento de plantas principalmente para avaliação de genótipos potenciais para produção. Este estudo incluiu 16 genótipos de cana-de-açúcar em 14 locais e avaliados por quatro anos, ou seja, quatro cortes (cana planta, primeira, segunda e terceira soca). O objetivo consiste na estimação de parâmetros genéticos via REML/BLUP, estimação da média harmônica da performance relativa do valor genotípico (MHPRVG) e seleção simultânea dos genótipos para produtividade, adaptabilidade e estabilidade a partir do atributo Tonelada de Pol por Hectare (TPH) avaliado em múltiplos ambientes na cultura da cana-de-açúcar. Os valores da herdabilidade média genotípica apresentaram-se altos, o que leva a altas acurácias na seleção genotípica (superiores a 90%). Os genótipos IACSP963076, IACSP963048 e IACSP985024, primeiro corte, IACSP963048, IACSP963076, IACSP962073 e IACSP985024, segundo corte, IACSP962073, IACSP962008 e IACSP963076, terceiro corte e IACSP963076, IACSP985024, IACSP962073 e IACSP962008, quarto corte, foram os selecionados quando se considera os atributos de estabilidade, adaptabilidade e produtividade (TPH), simultaneamente e, portanto, são os genótipos recomendados para o plantio nesta macro região estudada. Em termos de valores médios, os genótipos recomendados pela seleção simultânea baseada na MHPRVG são os genótipos IACSP963076, IACSP962073, IACSP985024 e IACSP962008, alcançando um valor médio de 2,8% acima da média geral dos quatro cortes. A metodologia da MHPRVG via modelos mistos é uma abordagem de grande utilidade na avaliação e capitalização da interação GxA quando se trata de um conjunto de dados desbalanceados.

Palavras-chave: adaptabilidade, estabilidade, Modelo misto, REML/BLUP

1. INTRODUÇÃO

A importância socioeconômica do cultivo da cana-de-açúcar (*Saccharum* spp.) é um fato observado desde os tempos coloniais da implantação desta cultura no Brasil. É a cultura mais utilizada na produção de açúcar para o consumo interno e produto de exportação, além de ser utilizada para a produção de álcool e biomassa. A cana-de-açúcar também tem demonstrado importância estratégica e ambiental, pela produção do biocombustível como fonte renovável de energia, menos poluente que combustíveis derivados do petróleo, e sua captação de carbono pela alta produção de biomassa. A área cultivada de cana-de-açúcar vem crescendo significativamente no Brasil, visto que, segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE), em 2003 plantou-se uma área de aproximadamente 5,38 milhões de hectares em comparação a 9,6

milhões de hectares em 2011 (UNICA,2013). O que leva aos melhoristas a buscarem genótipos que exibam maiores produtividades nas diferentes regiões de plantio.

Há um fator muito importante a ser considerado quando se trata de experimentos multi ambientais em estágios finais do melhoramento de plantas que é o fenômeno da interação genótipo por ambiente (GxA). Este fenômeno provém da interação de duas grandezas diferentes, inversamente proporcionais, das quais, para uma análise melhor, em termos de média, recomenda-se a utilização da média harmônica.

Segundo Resende (2004), no contexto dos modelos mistos, a seleção simultânea por produtividade, estabilidade e adaptabilidade pode ser realizada pelo método da média harmônica da performance relativa do valor genotípico (MHPRVG) predito.

Este trabalho objetiva a estimação de parâmetros genéticos via REML/BLUP, estimação da MHPRVG e seleção simultânea dos genótipos para produtividade, adaptabilidade e estabilidade a partir do atributo Tonelada de Pol por Hectare (TPH) avaliado em múltiplos ambientes na cultura da cana-de-açúcar.

2. MATERIAL E MÉTODOS

Encontra-se no CAPÍTULO 2 - MATERIAL E MÉTODOS, itens 1. MATERIAL e 2.2. Média Harmônica da Performance Relativa do Valor Genotípico (RESENDE, 2004).

3. RESULTADOS E DISCUSSÕES

3.1. Estimativa de componentes de variância, de parâmetros genéticos e Média Harmônica da Performance Relativa do Valor Genotípico (MHPRVG) via BLUP

3.1.1. Estimativa de componentes de variância e de parâmetros genéticos

Na Tabela 1 estão as estimativas dos componentes de variância e parâmetros genéticos para o atributo Tonelada de Pol por Hectare (TPH) para os quatro cortes avaliados, estimados pelo método REML.

Pode-se verificar pela Tabela 1 que as estimativas de herdabilidades individuais no

sentido amplo (h_g^2) têm um comportamento crescente do primeiro ao quarto corte. A herdabilidade média dos genótipos (h_{mg}^2) através dos ambientes e dos blocos foi maior no terceiro corte, entretanto alta também nos outros cortes. A raiz quadrada destes valores fornece a acurácia de seleção dos genótipos ($\hat{r}_{gg}$), a qual variou de 90,52 a 92,76% e é considerada, segundo Resende e Duarte (2007), como uma classe de precisão muito alta. Portanto, a experimentação realizada permite uma boa seleção para todos os locais simultaneamente em cada corte. A correlação genotípica entre as performances através dos ambientes ($\hat{r}_{gloc}$), apresentou-se de baixa a moderada, variando de 38 a 58%, o que indica uma coincidência de 38 a 58% na seleção de genótipos específicos para cada ambiente, ou seja, se fosse selecionar genótipos em cada ambiente no primeiro corte, a coincidência dessa seleção através dos ambientes seria de 38%.

Os coeficientes de determinação (b^2) ao longo de todos os cortes foram baixos e, desse modo, evidenciam a homogeneidade dos blocos nos experimentos. Já os coeficientes de determinação do efeito ambiental dentro do bloco (i^2) mostram o quanto da variação existente é devido a interação GxA, o que também justifica ser feita uma análise de interação GxA e sua capitalização na seleção de genótipos superiores.

A presença de uma variabilidade genética significativa é verificada pelo que demonstra as estimativas de herdabilidades individuais no sentido amplo dentro do bloco (h_g^2) e suas estimativas do erro padrão. Em comparação com a magnitude da h_g^2 , as estimativas do erro padrão foram de pequenas magnitudes, o que garante que a h_g^2 não passe zero no limite mais baixo do intervalo de confiança, caso contrário as h_g^2 não seriam significativas.

O coeficiente de variação genotípica ($CV_g\%$) variou de 7,6% a 11,9% entre genótipos analisados. Isto indica que há uma considerável variação genética e uma possibilidade de seleção efetiva entre genótipos, visto que sete dos 16 genótipos avaliados (44%) são variedades comerciais.

Os coeficientes de variação experimental foram de baixa magnitude, o que, também, expressa o bom controle experimental.

Houve um decréscimo na média de produtividade de TPH do primeiro ao quarto corte. Geralmente a produtividade das canas socas tende a diminuir gradativamente com o aumento dos estágios de corte, principalmente com a utilização de colheita mecanizada. Este fato é muito importante e pode direcionar, num futuro próximo, à seleção de genótipos de cana que sejam adaptados à colheita mecanizada.

Tabela 1. Estimativas de parâmetros genéticos para Tonelada de pol por hectare em genótipos de cana-de-açúcar nos quatro cortes.

Estimativas	Corte 1	Corte 2	Corte 3	Corte 4
$\hat{\sigma}_b^2$	0,5745	0,0018	0,1298	0,3183
$\hat{\sigma}_g^2$	2,2544	1,2893	1,3414	1,6224
$\hat{\sigma}_i^2$	2,1865	1,8352	1,2408	1,0210
$\hat{\sigma}_e^2$	5,3451	2,6437	1,5083	1,6588
$\hat{r}_{gloc}$	0,4924	0,5874	0,4805	0,3862
ρ	0,4841	0,5418	0,6426	0,6410
b^2	0,0555	0,0003	0,0308	0,0689
h_g^2	0,2176±0,04	0,2234±0,04	0,3178±0,05	0,3511±0,07
h_{mg}^2	0,8194	0,8503	0,8604	0,8306
$\hat{r}_{gg}$	0,9052	0,9221	0,9276	0,9114
i^2	0,2110	0,3181	0,2940	0,2210
$CV_g\%$	8,3105	7,5985	8,9651	11,8717
$CV_e\%$	12,7964	10,8808	9,5064	12,0041
μ	18,0672	14,9433	12,9189	10,7292

$\hat{\sigma}_b^2$: Variância do efeito de blocos; $\hat{\sigma}_g^2$: Variância genotípica; $\hat{\sigma}_i^2$: Variância da interação genótipo x ambiente; $\hat{\sigma}_e^2$: Variância residual; $\hat{r}_{gloc}$: Correlação genotípica entre as performances através dos ambientes; ρ : Repetibilidade individual dentro de ambiente; b^2 : Coeficiente de determinação do efeito de bloco; h_g^2 : herdabilidade individual no sentido amplo dentro do bloco; h_{mg}^2 : herdabilidade média genotípica; $\hat{r}_{gg}$: acurácia de seleção genotípica; i^2 : Coeficiente de determinação do efeito ambiental dentro do bloco; $CV_g\%$: Coeficiente de variação genotípica; $CV_e\%$: Coeficiente de variação experimental; μ : Média geral em toneladas de pol por hectare.

3.1.2. Seleção conjunta para produtividade, estabilidade e adaptabilidade via MHPRVG

Segundo Resende (2007) os valores genotípicos preditos para o genótipo i em cada local j usa simultaneamente os dados de todos os locais e são dados por VG_{ij} : $\mu_j + g_i + ge_{ij}$. Nesse caso, tanto o g quanto o ge são preditos com maior precisão pois todo o conjunto de dados é usado bem como os ruídos da interação são eliminados quando se produzem os BLUP's de ge .

Nas Tabelas 2, 3, 4 e 5 estão apresentados os resultados sobre a estabilidade (MHVG), adaptabilidade (PRVG) e estabilidade e adaptabilidade simultaneamente (MHPRVG), para o atributo TPH dos nove genótipos e sete variedades comerciais de cana-de-açúcar avaliados nos respectivos locais em quatro cortes.

Segundo Resende (2007), quanto mais baixo o erro padrão da performance

genotípica através dos locais maior é a MHVG implicando na seleção para produtividade e estabilidade. A diferença neste método para a análise de estabilidade pode residir na acurácia de estimação dos efeitos envolvidos no cálculo do VG_{ij} e também da própria utilização de tal valor na análise de estabilidade. Ao contrário do que acontece em outros métodos onde não se considera o efeito dos genótipos como um fator aleatório e consequentemente não é utilizado.

A Tabela 2 está ordenada pela MHPRVG, percebe então que existe diferenças no ordenamento quando se considera somente a estabilidade e produtividade (MHVG) na seleção de genótipos. Pelos resultados obtidos observa-se que somente variedades comerciais ficaram acima da média geral (μ), dos quais foram: RB92579 (19,01), SP813250 (18,60) e CTC2 (18,26) para MHVG. Um dos genótipos testados ficou próximo da μ e pode ser considerado para a seleção: IACSP963076 (17,85).

Segundo Resende (2004), adaptabilidade é a capacidade do genótipo de melhorar em resposta à melhoria do ambiente. Em termos de adaptabilidade, os valores genotípicos preditos (ou os dados originais) são expressos como proporção da média geral de cada local (μ_j) e, posteriormente, obtém-se o valor médio desta proporção através dos locais. Pela multiplicação dos valores da PRVG pela média geral (μ) obtém-se já os valores na escala do atributo mensurado. Neste caso, já pode-se inferir sobre os genótipos mais adaptados pela coluna de $PRVG \cdot \mu$, onde os genótipos mais adaptados, desconsiderando-se as variedades comerciais, são: O IACSP963076 (19,57), IACSP963048 (18,72) e IACSP985024 (18,16). Pela PRVG seleciona-se genótipos responsivos à melhoria do ambiente.

Com o objetivo de selecionar os genótipos a partir da produtividade, adaptabilidade e estabilidade deve-se procurar aqueles cuja estimativa da MHPRVG seja igual ou acima de 1. Dessa forma, para o primeiro corte (Tabela 2), houve três genótipos que se enquadram neste parâmetro: IACSP963076, IACSP963048 e IACSP985024. Os seus respectivos valores de MHPRVG indicam a superioridade média do genótipo em relação à média do ambiente em que for cultivado, ou seja, segundo Resende (2004) é uma inferência sobre a produtividade esperada, tais estimativas de MHPRVG podem ser usadas quando for considerar o plantio em vários locais com variados padrões de interação GxA.

Dos três genótipos selecionados com base na MHPRVG, o IACSP985024 mostrou-se competitivo porém não apresentou ganho sobre a média geral dos ambientes em que

foi testado (MHPRVG = 1). Já a seleção dos genótipos IACSP963076 e IACSP963048 propicia um ganho de 5,5% sobre a média geral dos 13 ambientes, considerando simultaneamente a produtividade, estabilidade e adaptabilidade através dos locais.

Corroborar-se com trabalhos de Sturion e Resende (2005), Oliveira et al. (2005) e Carbonell et al. (2007) de que o método da MHPRVG apresenta a vantagem de fornecer resultados na própria escala e medição do caráter avaliado, os quais podem ser interpretados diretamente como valores genéticos, ou seja, a estimativa da MHPRVG multiplicada pela média geral (μ) fornece a própria média predita do genótipo quando plantado em outro ambiente, valor já penalizado pela instabilidade e capitalizado pela adaptabilidade. Neste caso, para os dois melhores genótipos, IACSP963076 e IACSP963048, as médias seriam 19,50 e 18,66 toneladas de pol por hectare, respectivamente, ou seja, 8 e 3% acima da média de todos os genótipos analisados.

Tabela 2. Estabilidade de valores genotípicos (MHVG), adaptabilidade de valores genotípicos (PRVG), ordenada pela estabilidade e adaptabilidade de valores genotípicos (MHPRVG) para Tonelada de Pol por Hectare de 16 genótipos de cana-de-açúcar no primeiro corte.

Genótipo	MHVG	PRVG	PRVG* μ	MHPRVG	MHPRVG* μ
CTC2**	18,26	1,14	20,61	1,14	20,57
SP813250**	18,60	1,12	20,18	1,11	20,04
IACSP963076	17,85	1,08	19,57	1,08	19,50
RB867515**	17,52	1,07	19,30	1,06	19,23
RB92579**	19,01	1,07	19,28	1,06	19,20
IACSP963048	16,81	1,04	18,72	1,03	18,66
CTC4**	15,60	1,02	18,48	1,02	18,42
IACSP985024	16,56	1,01	18,16	1,00	18,07
IACSP933046**	15,95	0,98	17,72	0,98	17,69
IACSP962073	16,15	0,98	17,71	0,98	17,66
IACSP962008	15,89	0,97	17,51	0,97	17,48
IACSP986202	15,81	0,97	17,50	0,97	17,44
RB72454**	15,52	0,97	17,53	0,96	17,43
IACSP985046	15,08	0,93	16,79	0,93	16,73
IACSP952289	13,40	0,85	15,34	0,84	15,19
IACSP972084	12,91	0,81	14,70	0,81	14,60

μ : média geral = 18,0672 toneladas de pol por hectare; ** Variedade comercial.

Na Tabela 3 estão os resultados das análises de interação GxA para o segundo corte. Os genótipos estão ordenados pela MHPRVG.

Considerando a produtividade e estabilidade dos genótipos, o IACSP963076 é selecionado também no segundo corte como genótipo produtivo e estável. Neste segundo

corte o genótipo IACSP963048 também destaca-se pela sua estabilidade e produtividade acima da média geral.

Mas no contexto deste método, pretende-se obter genótipos considerando os três atributos simultaneamente, via MHPRVG. Sendo assim, os genótipos selecionados, desconsiderando as variedades comerciais, foram: IACSP963048, IACSP963076, IACSP962073 e IACSP985024. Os genótipos IACSP963048 e IACSP963076 foram também selecionados para o primeiro corte.

Se os quatro genótipos selecionados fossem plantados na rede experimental ou em ambientes desconhecidos mas dentro da macro região testada, visto que tanto o efeitos de genótipos e de interação GxA foram considerados aleatórios, obter-se-ia uma média de 2,25% a mais do que a média geral no segundo corte ($\mu = 14,94$), ou seja, uma nova média de 15,28 toneladas de pol por hectare.

Tabela 3. Estabilidade de valores genotípicos (MHVG), adaptabilidade de valores genotípicos (PRVG), ordenada pela estabilidade e adaptabilidade de valores genotípicos (MHPRVG) para Tonelada de Pol por Hectare de 16 genótipos de cana-de-açúcar no segundo corte.

Genótipo	MHVG	PRVG	PRVG* μ	MHPRVG	MHPRVG* μ
CTC2**	15,77	1,15	17,16	1,14	17,05
CTC4**	15,06	1,08	16,17	1,08	16,15
RB867515**	15,18	1,05	15,70	1,05	15,62
IACSP963048	14,98	1,05	15,71	1,04	15,60
IACSP963076	14,76	1,04	15,53	1,03	15,43
RB92579**	14,99	1,03	15,44	1,03	15,36
SP813250**	14,62	1,03	15,33	1,02	15,31
IACSP933046**	14,00	1,02	15,20	1,02	15,17
IACSP962073	14,47	1,01	15,12	1,01	15,11
IACSP985024	14,45	1,01	15,15	1,01	15,09
IACSP962008	14,16	0,99	14,75	0,98	14,71
IACSP985046	13,68	0,96	14,32	0,95	14,25
IACSP952289	13,71	0,95	14,23	0,95	14,19
RB72454**	13,23	0,93	13,91	0,92	13,80
IACSP986202	12,76	0,90	13,41	0,89	13,30
IACSP972084	11,63	0,81	12,17	0,81	12,12

μ : média geral = 14,9433 toneladas de pol por hectare; **Variedade comercial.

Na Tabela 4 estão os resultados das análises de interação GxA para o terceiro corte. Os genótipos estão ordenados pela MHPRVG.

Considerando somente a produtividade e estabilidade dos genótipos, nenhum genótipo que não seja variedade comercial se destacou. Todos os genótipos testados

foram abaixo da média geral quando foi feita a análise somente pela produtividade e estabilidade. Um genótipo que se aproxima da média geral e que pode ser levado em consideração é o IACSP962073, que teve sua MHVG de 12,30, sendo 4,8% abaixo da média geral.

Quando se considera os três atributos simultaneamente, dentre os 8 genótipos que estão acima da média, três deles não são variedades comerciais. Na seleção destes três genótipos (IACSP962073, IACSP962008 e IACSP963076), a nova média da população seria 3,33% acima da média geral de todos os genótipos testados. Observa-se o genótipo IACSP963076 novamente na seleção entre os melhores genótipos dos quais fora selecionado também no primeiro e segundo corte, ou seja, um genótipo que demonstra uma alta produtividade, adaptabilidade e estabilidade até o terceiro corte.

Tabela 4. Estabilidade de valores genotípicos (MHVG), adaptabilidade de valores genotípicos (PRVG), ordenada pela estabilidade e adaptabilidade de valores genotípicos (MHPRVG) para Tonelada de Pol por Hectare de 16 genótipos de cana-de-açúcar no terceiro corte.

Genótipo	MHVG	PRVG	PRVG* μ	MHPRVG	MHPRVG* μ
RB867515**	13,91	1,18	15,19	1,16	14,99
CTC4**	12,26	1,14	14,68	1,13	14,61
RB92579**	13,74	1,10	14,17	1,09	14,13
CTC2**	11,53	1,09	14,14	1,09	14,03
IACSP933046**	11,80	1,08	13,92	1,08	13,89
IACSP962073	12,30	1,06	13,66	1,05	13,60
IACSP962008	12,17	1,04	13,42	1,03	13,34
IACSP963076	11,53	1,03	13,26	1,02	13,13
IACSP985024	11,22	0,98	12,71	0,98	12,65
SP813250**	10,72	0,97	12,59	0,97	12,48
RB72454**	11,13	0,97	12,49	0,96	12,46
IACSP952289	10,38	0,91	11,71	0,90	11,68
IACSP985046	10,01	0,89	11,52	0,89	11,45
IACSP963048	9,78	0,88	11,34	0,87	11,22
IACSP972084	10,02	0,87	11,27	0,86	11,10
IACSP986202	9,00	0,84	10,90	0,82	10,64

μ : média geral = 12,9189 toneladas de pol por hectare; **Variedade comercial.

Na Tabela 5 estão os resultados das análises de interação GxA para o quarto corte, sendo que os genótipos estão ordenados pela MHPRVG.

Observa-se que dos 4 genótipos que possuem produtividade acima da média geral e são estáveis em manter essa produtividade, três deles são novos genótipos testados. Sendo eles: IACSP963076, IACSP985024 e o IACSP962073, que estão em média 2,65%

acima da média geral de todos os genótipos testados.

Para o quarto corte, dos nove genótipos que estão com a MHPRVG acima da média geral, ou seja, genótipos que geram um ganho genotípico em produtividade, quatro deles são oriundos dos novos genótipos testados, sendo eles: IACSP963076 (12,06), IACSP985024 (11,68), IACSP962073 (11,51) e IACSP962008 (11,01). Quando os quatro forem selecionados, a estimativa do ganho em produtividade será, em média, de 7,75% acima da média geral de todos os genótipos. Dentre estes quatro selecionados destaca-se o IACSP963076. Este genótipo mostrou-se produtivo, adaptável e estável simultaneamente ao longo dos quatro cortes analisados.

Tabela 5. Estabilidade de valores genotípicos (MHVG), adaptabilidade de valores genotípicos (PRVG), ordenada pela estabilidade e adaptabilidade de valores genotípicos (MHPRVG) para Tonelada de Pol por Hectare de 16 genótipos de cana-de-açúcar no quarto corte.

Genótipo	MHVG	PRVG	PRVG* μ	MHPRVG	MHPRVG* μ
CTC2**	12,86	1,27	13,63	1,26	13,49
IACSP963076	11,29	1,13	12,12	1,12	12,06
CTC4**	10,26	1,13	12,14	1,12	12,04
IACSP985024	10,95	1,09	11,69	1,09	11,68
IACSP962073	10,80	1,08	11,54	1,07	11,51
IACSP933046**	10,35	1,05	11,31	1,05	11,29
RB867515**	10,61	1,05	11,22	1,04	11,19
RB92579**	9,56	1,03	11,09	1,03	11,08
IACSP962008	10,45	1,03	11,04	1,03	11,01
SP813250**	9,92	0,98	10,53	0,98	10,50
IACSP963048	9,40	0,93	10,03	0,93	10,01
IACSP985046	9,09	0,92	9,87	0,91	9,78
IACSP952289	8,55	0,86	9,25	0,86	9,18
IACSP972084	8,29	0,83	8,95	0,83	8,87
IACSP986202	8,03	0,83	8,93	0,82	8,80
RB72454**	7,78	0,81	8,66	0,79	8,44

μ : média geral = 10,7292 toneladas de pol por hectare; **Variedade comercial.

A Tabela 6 refere-se à média da MHVG, PRVG, MHPRVG e de produtividade nos quatro cortes analisados.

Em termos de média dos quatro cortes, observa-se que somente os genótipos IACSP963076, IACSP962073, IACSP985024 e IACSP962008 obtiveram uma MHPRVG acima de 1. Assim, se esses quatro genótipos forem plantados, há uma expectativa de produtividade de 2,8% acima da média geral dos quatro cortes. Isto representa uma produtividade em torno de 14,6 toneladas de pol por hectare, em média.

Tabela 6. Valores médios dos quatro cortes para a estabilidade de valores genotípicos (MHVG), adaptabilidade de valores genotípicos (PRVG), ordenada pela estabilidade e adaptabilidade de valores genotípicos (MHPRVG) para Tonelada de Pol por Hectare de 16 genótipos de cana-de-açúcar.

Genótipo	MHVG	PRVG	PRVG* μ	MHPRVG	MHPRVG* μ
CTC2**	14,6050	1,1625	16,4664	1,1575	16,3956
CTC4**	13,2950	1,0925	15,4749	1,0875	15,4041
RB867515**	14,3050	1,0875	15,4041	1,0775	15,2624
IACSP963076	13,8575	1,0700	15,1562	1,0625	15,0499
RB92579**	14,3250	1,0575	14,9791	1,0525	14,9083
SP813250**	13,4650	1,0250	14,5188	1,0200	14,4479
IACSP933046**	13,0250	1,0325	14,6250	1,0325	14,6250
IACSP962073	13,4300	1,0325	14,6250	1,0275	14,5542
IACSP985024	13,2950	1,0225	14,4834	1,0200	14,4479
IACSP962008	13,1675	1,0075	14,2709	1,0025	14,2001
IACSP963048	12,7425	0,9750	13,8105	0,9675	13,7043
IACSP985046	11,9650	0,9250	13,1023	0,9200	13,0315
RB72454**	11,9150	0,9200	13,0315	0,9075	12,8544
IACSP952289	11,5100	0,8925	12,6420	0,8875	12,5711
IACSP986202	11,4000	0,8850	12,5357	0,8750	12,3941
IACSP972084	10,7125	0,8300	11,7567	0,8275	11,7212

μ : média geral dos quatro cortes = 14,1646 toneladas de pol por hectare; **Variedade comercial.

4. CONCLUSÕES

- Os valores da herdabilidade média genotípica apresentaram-se altos, o que leva a altas acurácias na seleção genotípica (superiores a 90%)
- Os genótipos IACSP963076, IACSP963048 e IACSP985024, primeiro corte, IACSP963048, IACSP963076, IACSP962073 e IACSP985024, segundo corte, IACSP962073, IACSP962008 e IACSP963076, terceiro corte e IACSP963076, IACSP985024, IACSP962073 e IACSP962008, quarto corte, foram os selecionados quando se considera os atributos de estabilidade, adaptabilidade e produtividade (TPH), simultaneamente e, portanto, são os genótipos recomendados para o plantio na região centro-sul nos estados de SP, GO, MT e MG dos quais foram estudados.
- Em termos de valores médios, os genótipos recomendados pela seleção simultânea baseada na MHPRVG são os genótipos IACSP963076, IACSP962073, IACSP985024 e IACSP962008, alcançando um valor médio de 2,8% acima da

média geral dos quatro cortes.

- A metodologia da MHPRVG via modelos mistos mostra-se uma abordagem de grande utilidade na avaliação e capitalização da interação GxA quando se trata de um conjunto de dados desbalanceados.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ANNICCHIARICO, P. Cultivar adaptation and recommendation from alfalfa trials in Northern Italy. **Journal of Genetics & Breeding**, Berlin, v. 46, n. 1, p. 269-278, 1992.
- CARBONELL, S. A. M.; CHIORATO, A. F.; RESENDE, M. D. V.; DIAS, L. A. S.; BERALDO, A. L. A.; PERINA, E. F. Estabilidade de cultivares e linhagens de feijoeiro em diferentes ambientes no estado de São Paulo. **Bragantia**, Campinas, v. 66, n. 2, p.193-201, 2007.
- OLIVEIRA, R.; RESENDE, M. D. V. ; DAROS, E. ; BESPALHOK FILHO, J. C.; ZAMBON, J. L.; IDO, O.; KOEHLER, Henrique . Genotypic evaluation and selection of sugarcane clones in three environments in the state of Paraná. **Crop breeding and applied biotechnology**, Viçosa, v. 5, n. 4, p. 426-434, 2005.
- RESENDE, M. D. V. **Métodos Estatísticos Ótimos na Análise de Experimentos de Campo**. Colombo: Embrapa Florestas, Ed. 1, 2004, 57 p. (Documentos, 100).
- RESENDE, M.D.V. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas. 2007. 561p.
- RESENDE, M.D.V. de; DUARTE, J.B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v.37, p.182-194, 2007.
- SAS Institute Inc. 2010. SAS/STAT® 9.22 User's Guide. Cary, NC: SAS Institute Inc.
- STURION, J. A.; RESENDE, M.D.V. Seleção de progênies de erva-mate (*Ilex paraguariensis* St. Hil.) para produtividade, estabilidade e adaptabilidade temporal de massa foliar. **Boletim de pesquisa florestal**, Colombo, n. 50, p. 37-51, 2005.
- UNICA, União da Indústria de cana-de-açúcar, Fonte: www.unicadata.com.br, Acesso: 05 set. 2013.
- WRIGHT, J.W.; PAULEY, S.S.; POLK, R.B.; JOKELA, J.J. Performance of Scotch pine varieties in North Central Region, **Silvae Genetica**, v. 15, 0. 101-110, 1966.

CAPÍTULO 5 - Considerações finais e conclusões

1. CONSIDERAÇÕES FINAIS

Os três métodos aplicados à análise da interação GxA levam em consideração diferentes abordagens.

Assim, o método de Eberhart and Russell (E-R) (1966) procura captar a interação GxA em suas análises utilizando um conceito em que todos os efeitos são fixos e a regressão é baseada em um Índice Ambiental calculado com os resultados de uma análise previa do modelo e inseridos como variável explicativa. Já na análise de MHPRVG (RESENDE, 2004), o fator ambiente (no caso, usina) é considerado fixo e o fator genótipo é considerado aleatório, dessa forma é estimado os valores de média harmônica e performance relativa conforme procedimentos citados pelo autor. Para o método proposto por Su et al. (2006), norma de reação (NR), numa abordagem bayesiana, todos os fatores são considerados aleatórios e, portanto, cada um deles, como um fator aleatório, possui distribuição de densidade de probabilidade a priori. Após a realização da estatística, obtém-se as estimativas dos parâmetros genéticos, assim como os respectivos coeficientes de regressão e intercepto. A variável explicativa, que no caso é o efeito do ambiente, é estimada simultaneamente no modelo. Entre as regressões envolvidas, E-R estima coeficientes de regressão que são preditores de valores fenotípicos enquanto que na NR os coeficientes de regressão estimados para cada genótipo são preditores de valores genotípicos, ou seja, livres de efeitos residuais.

Tanto na NR quanto em E-R são estimados índices ambientais (I_j), assim o genótipo é regredido utilizando-se esses valores de I_j como variáveis explicativas. Ambos os métodos apresentaram I_j semelhantes embora quando houve desbalanceamento, ou seja, o genótipo i não estava presente em um dos ambientes j , não foi possível estimar pela metodologia de E-R a resposta para aquele ambiente em que o genótipo i não estava presente. É interessante observar que na abordagem de Su et al. (2006), por construção, é suposta uma distribuição binormal entre os índices ambientais e a resposta (produção). Assim, como bem demonstrado, por exemplo em Hogg and Craig, (1965, pág. 103-104) a reta sempre existe pois a esperança de Y dado X é uma reta, ou seja, $E(Y | X) = g + bX$.

De maneira geral, pode-se elencar na Tabela 7 como é inferida a estabilidade e

adaptabilidade pelos métodos abordados nesta Tese.

Tabela 7. Resumo de estimação e interpretação da interação GxA (adaptabilidade e estabilidade) pelos métodos de Eberhart and Russell (E-R), Resende (MHPRVG) e Su et al. (NR).

Método	Efeitos fixos	Efeitos aleatórios	Modelo estatístico	Estabilidade	Adaptabilidade favorável
E-R	a, b, g	e	$y_{ijk} = \mu + g_i + b_i I_j + D_i + \beta_k(I_j) + e_{ijk}$	$MS-D_i \cong 0$ $b_i \cong 0$;	$MS-D_i \cong 0$ $b_i \geq 1$
MHPRVG	a	g, b, i, e	$y = Xr + Tb + Zg + Wi + e$	Maior valor da MHVG _i *	$PRVG_i > 1$
NR	-	a, b, g, f, e	$y = X\beta + Wh + Z_{\alpha}g + Z_i b + e$	$b_i \cong 0$	$b_i > 0$

MHPRVG: média harmônica da performance relativa do valor genotípico; NR: norma de reação com covariáveis desconhecidas; a: ambiente, b: bloco; g: genótipo; i: interação GxA; f: efeito ambiental; e: erro, MS-D_i: Quadrado médio do desvio do genótipo i; b_i: Coeficiente de regressão; *Estabilidade e produtividade simultaneamente.

Na Tabela 8 está o ordenamento dos genótipos que possuem alta produtividade, estabilidade e adaptabilidade a partir dos três métodos abordados para o primeiro corte em ensaios multi ambientais de cana-de-açúcar.

Tabela 8. Ordenamento dos genótipos de acordo com o método de análise para estabilidade e adaptabilidade para o primeiro corte em ensaios multi ambientais de cana-de-açúcar.

Genótipo	E-R	NR	MHPRVG
	Ordenamento		
CTC2 ^{VP}	2	1	1
CTC4 ^{VP}	13	7	7
IACSP933046 ^{VP}	12	13	9
IACSP952289	15*	15	15
IACSP962008	10	11	11
IACSP962073	8	10	10
IACSP963048	6	6	6
IACSP963076	4	3	3
IACSP972084	16	16	16
IACSP985024	7*	8	8
IACSP985046	14	14	14
IACSP986202	11*	12	12
RB72454 ^{VP}	9	9	13
RB867515 ^{VP}	5	4	4
RB92579 ^{VP}	1*	5	5
SP813250 ^{VP}	3*	2	2

E-R: Regressão linear fenotípica conjunta; NR: Norma de reação com covariáveis desconhecidas; MHPRVG: Média harmônica da performance relativa do valor genotípico; VP: Variedade padrão; *: Adaptabilidade preferencial.

Na Tabela 9 está o ordenamento dos genótipos que possuem estabilidade e adaptabilidade a partir dos três métodos abordados para o segundo corte em ensaios multi ambientais de cana-de-açúcar.

Tabela 9. Ordenamento dos genótipos de acordo com o método de análise para estabilidade e adaptabilidade para o segundo corte em ensaios multi ambientais de cana-de-açúcar.

Genótipo	E-R	NR	MHPRVG
	Ordenamento		
CTC2 ^{VP}	1*	1	1
CTC4 ^{VP}	2	2	2
IACSP933046 ^{VP}	11	9	8
IACSP952289	13	14	13
IACSP962008	10	11	11
IACSP962073	7	8	9
IACSP963048	5*	5	4
IACSP963076	3*	3	5
IACSP972084	16*	16	16
IACSP985024	9	10	10
IACSP985046	12*	12	12
IACSP986202	15*	15	15
RB72454 ^{VP}	14*	13	14
RB867515 ^{VP}	4	6	3
RB92579 ^{VP}	6*	4	6
SP813250 ^{VP}	8	7	7

E-R: Regressão linear fenotípica conjunta; NR: Norma de reação com covariáveis desconhecidas; MHPRVG: Média harmônica da performance relativa do valor genotípico; VP: Variedade padrão; *: Adaptabilidade preferencial.

Na Tabela 10 está o ordenamento dos genótipos que possuem estabilidade e adaptabilidade a partir dos três métodos abordados para o terceiro corte em ensaios multi ambientais de cana-de-açúcar.

Tabela 10. Ordenamento dos genótipos de acordo com o método de análise para estabilidade e adaptabilidade para o terceiro corte em ensaios multi ambientais de cana-de-açúcar.

Genótipo	E-R	NR	MHPRVG
	Ordenamento		
CTC2 ^{VP}	8*	3	4
CTC4 ^{VP}	3*	2	2
IACSP933046 ^{VP}	5*	5	5
IACSP952289	12	12	12
IACSP962008	7	8	7
IACSP962073	4	6	6
IACSP963048	14	14	14
IACSP963076	6*	7	8
IACSP972084	16*	16	15
IACSP985024	11	10	9
IACSP985046	13	13	13
IACSP986202	15	15	16
RB72454 ^{VP}	10	11	11
RB867515 ^{VP}	1	1	1
RB92579 ^{VP}	2	4	3
SP813250 ^{VP}	9	9	10

E-R: Regressão linear fenotípica conjunta; NR: Norma de reação com covariáveis desconhecidas; MHPRVG: Média harmônica da performance relativa do valor genotípico; VP: Variedade padrão; *: Adaptabilidade preferencial.

Na Tabela 11 está o ordenamento dos genótipos que possuem estabilidade e adaptabilidade a partir dos três métodos abordados para o primeiro corte em ensaios multi ambientais de cana-de-açúcar.

Tabela 11. Ordenamento dos genótipos de acordo com o método de análise para estabilidade e adaptabilidade para o quarto corte em ensaios multi ambientais de cana-de-açúcar.

Genótipo	E-R	NR	MHPRVG
	Ordenamento		
CTC2 ^{VP}	1	1	1
CTC4 ^{VP}	8	3	3
IACSP933046 ^{VP}	7	8	6
IACSP952289	13*	13	13
IACSP962008	6	9	9
IACSP962073	3	4	5
IACSP963048	11	11	11
IACSP963076	2*	2	2
IACSP972084	15*	15	14
IACSP985024	4	5	4
IACSP985046	12	12	12
IACSP986202	14	14	15
RB72454 ^{VP}	16*	16	16
RB867515 ^{VP}	5	6	7
RB92579 ^{VP}	9	7	8
SP813250 ^{VP}	10	10	10

E-R: Regressão linear fenotípica conjunta; NR: Norma de reação com covariáveis desconhecidas; MHPRVG: Média harmônica da performance relativa do valor genotípico; VP: Variedade padrão; *: Adaptabilidade preferencial.

Analisando os métodos abordados pelas Tabelas 8, 9, 10 e 11, se fossemos selecionar 25% dos genótipos, ou seja, neste caso, os quatro melhores, as metodologias coincidiriam 100% entre a norma de reação (NR) e média harmônica da performance relativa do valor genotípico (MHPRVG) para o primeiro corte, 50% para o segundo, 100% para o terceiro corte e 75% para o quarto corte.

Uma vez que, além de se identificar genótipos com adaptabilidade preferencial pela análise de regressão linear fenotípica conjunta, identifica-se também genótipos com maiores valores genotípicos e seus respectivos coeficientes de regressão da NR a fim de ditar o comportamento de tal genótipo pelo gradiente ambiental e, por fim, a predição de produção desses genótipos a partir da MHPRVG via BLUP obtendo-se simultaneamente genótipos produtivos, estáveis e adaptáveis. Pode-se destacar que os três métodos abordados foram complementares.

2. CONCLUSÕES

- Todos os métodos analisados nesta tese proporcionaram bons resultados e podem ser utilizados na avaliação da interação genótipo x ambiente (GxA), adaptabilidade e estabilidade;
- O método de Eberhart and Russell pode ser utilizado, a princípio, para uma visão geral do comportamento dos genótipos nos ambientes visto que fornece estimativas de coeficientes de regressão de valores fenotípicos;
- Os métodos de Eberhart and Russell e Norma de reação proporcionam uma melhor visualização gráfica da interação GxA, adaptabilidade e estabilidade;
- Os métodos da Norma de reação e MHPRVG possibilitam o uso de dados desbalanceados, utilizam o fator genótipo como aleatório, o que proporciona a estimação do valor genotípico do indivíduo e aumento na acurácia de seleção genotípica;
- A estimativa dos Índices Ambientais (I_j) foram semelhantes para os métodos de E-R e Norma de reação, para este conjunto de dados analisados;
- O método da MHPRVG fornece resultados na própria escala de medição do caráter avaliado, podendo ser interpretados diretamente como valores genotípicos para produtividade, estabilidade e adaptabilidade simultaneamente;
- O método da Norma de reação estima coeficientes de regressão que predizem valores genotípicos, o que permite aos melhoristas uma melhor exploração do valor genotípico do indivíduo no programa de melhoramento genético.

3. REFERÊNCIAS BIBLIOGRÁFICAS

- EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, n. 1, p. 36-40, 1966.
- HOGG, R.V.; CRAIG, A.T. **Introduction to mathematical statistics**, 2^a ed., New York, Macmillan Co., 1965. 388 p.
- RESENDE, M. D. V. **Métodos Estatísticos Ótimos na Análise de Experimentos de Campo**. Colombo: Embrapa Florestas, Ed. 1, 2004, 57 p. (Documentos, 100).
- SU, G.; MADSEN, P.; LUND, M. S.; D, SORENSEN; KORSGAARD, I.R.; JENSEN, J. Bayesian analysis of the linear reaction norm model with unknown covariates. **Journal of Animal Science**, v. 84, p. 1651–1657, 2006.

APÊNDICE

1. Diagnósticos de convergência

Antes de ser feita qualquer inferência a posteriori é essencial que se examine a convergência da cadeia de Markov de todos os parâmetros. Não se pode fazer uma inferência válida se a cadeia de Markov não convergiu. Uma ferramenta de diagnóstico muito efetiva é o gráfico de traço (trace plot), por exemplo. Entretanto iremos demonstrar mais algumas. As análises de convergência foram realizadas pelo software R (R DEVELOPMENT CORE TEAM, 2012) a partir do pacote *boa* (SMITH, 2007).

Para cada distribuição a posteriori, o pacote *boa* também reporta um resumo da estatística (**média a posteriori e desvios padrão** Tabela 12 do Apêndice), assim como **as autocorrelações** nas defasagens (lags) selecionadas e o **diagnóstico de Geweke** (todos estes são diagnósticos de convergência que ajudam na determinação da convergência da cadeia dos parâmetros do modelo).

A partir do pacote *boa* foi possível obter gráficos nos quais adicionam na checagem do diagnóstico da convergência. Com o **gráfico de traço**, há dois aspectos importantes para se examinar. Primeiro, deve-se checar se a média da cadeia de Markov estabilizou e aparenta constante no gráfico. Segundo, deve-se checar se a cadeia tem uma boa mistura e está "densa", no sentido que ela rapidamente percorra o suporte da distribuição para explorar ambas as pontas e as áreas da moda eficientemente. Depois, deve-se examinar o **gráfico de autocorrelação**, que indica o grau de autocorrelação para cada uma das amostras a posteriori. Altas correlações geralmente sugerem uma baixa mistura "Slow mixing". Finalmente, o gráfico da **densidade do "kernel"** que estima a distribuição marginal a posteriori para cada parâmetro.

Observa-se que para os quatro cortes analisados a média da cadeia de Markov estabilizou e aparenta constante no gráfico assim como tem uma boa mistura e encontra-se densa (Figura 1, 4, 7 e 10). As autocorrelações das amostras a posteriori reduzem rapidamente e tornam-se quase não existente depois do lag 5, isso pode ser verificado para todos os quatro cortes analisados (Figura 2, 5, 8 e 11). O diagnóstico de Geweke indica que nenhum parâmetro falhou no teste, isso deve-se às probabilidades significativamente altas encontradas para o Z-Score (Tabela 13).

Na Tabela 12 estão as estimativas médias a posteriori do amostrador de Gibbs de vários parâmetros genéticos, incluindo os elementos da matriz de covariância tais como

σ_g^2 , σ_b^2 e σ_{gb}^2 . A distribuição a posteriori da variância do coeficiente de regressão σ_b^2 foi significativamente afastada de zero para todas os quatro cortes avaliados, o que dá suporte à presença da interação GxE.

Tabela 12. Estimativa via norma de reação da média a posteriori e desvio padrão dos componentes de variâncias e covariância para o atributo Tonelada de Pol por Hectare em ensaios multi ambientais de cana-de-açúcar para quatro cortes.

Estimativas	Primeiro corte	Segundo corte	Terceiro corte	Quarto corte
σ_g^2	3,5222±1,8293	2,1681±1,1701	2,4169±1,2564	2,7035±1,4279
σ_b^2	0,0221± 0,0149	0,0413±0,0278	0,0248±0,0151	0,0155±0,0143
σ_{gb}	0,1973±0,1316	0,2644±0,1572	0,0749±0,0987	0,0663±0,1038
σ_h^2	45,0486±24,3610	13,4094±7,8219	29,9992±21,5904	26,5494±44,5048
σ_p^2	0,5274±0,3251	0,0924±0,0942	0,2181±0,1545	0,3119±0,2563
σ_e^2	7,6118±0,4646	4,5554±0,2928	2,6491±0,1933	2,7654±0,2448

Variância genotípica (σ_g^2), Variância do coeficiente de regressão genotípico (σ_b^2), Covariância genotípica e coeficiente de regressão da norma de reação (σ_{gb}), Variância do efeito do ambiente (σ_h^2), Variância do efeito de bloco (σ_p^2) e Variância do efeito residual (σ_e^2).

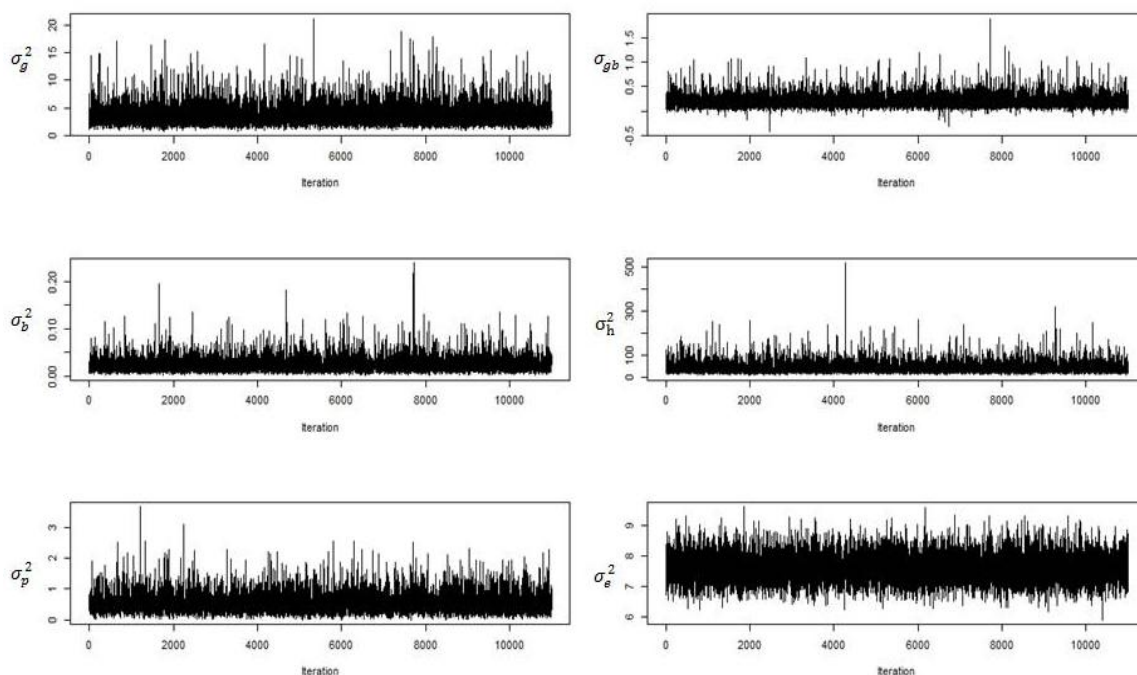


Figura 1. Gráfico de traço dos componentes de variância e co-variância estimados vs.

iterações de cadeia de Markov para o **primeiro corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

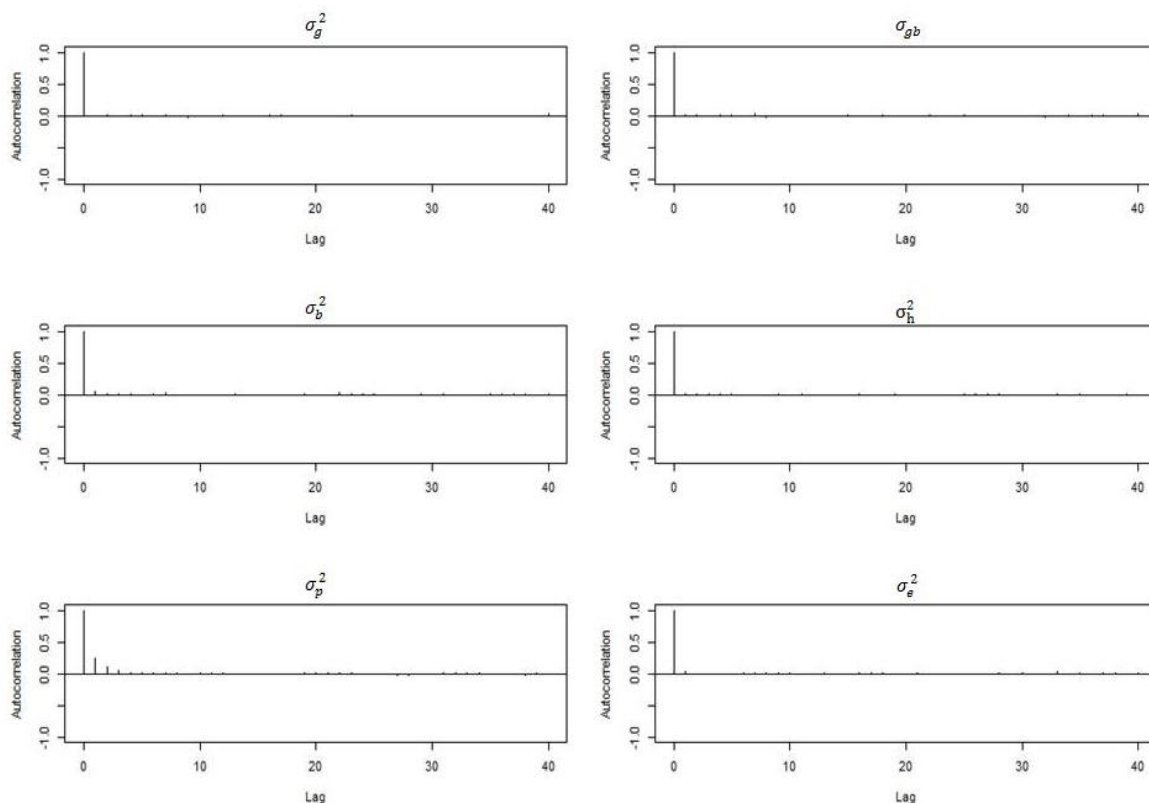


Figura 2. Gráfico das autocorrelações dos componentes de variância e co-variância estimados para o **primeiro corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

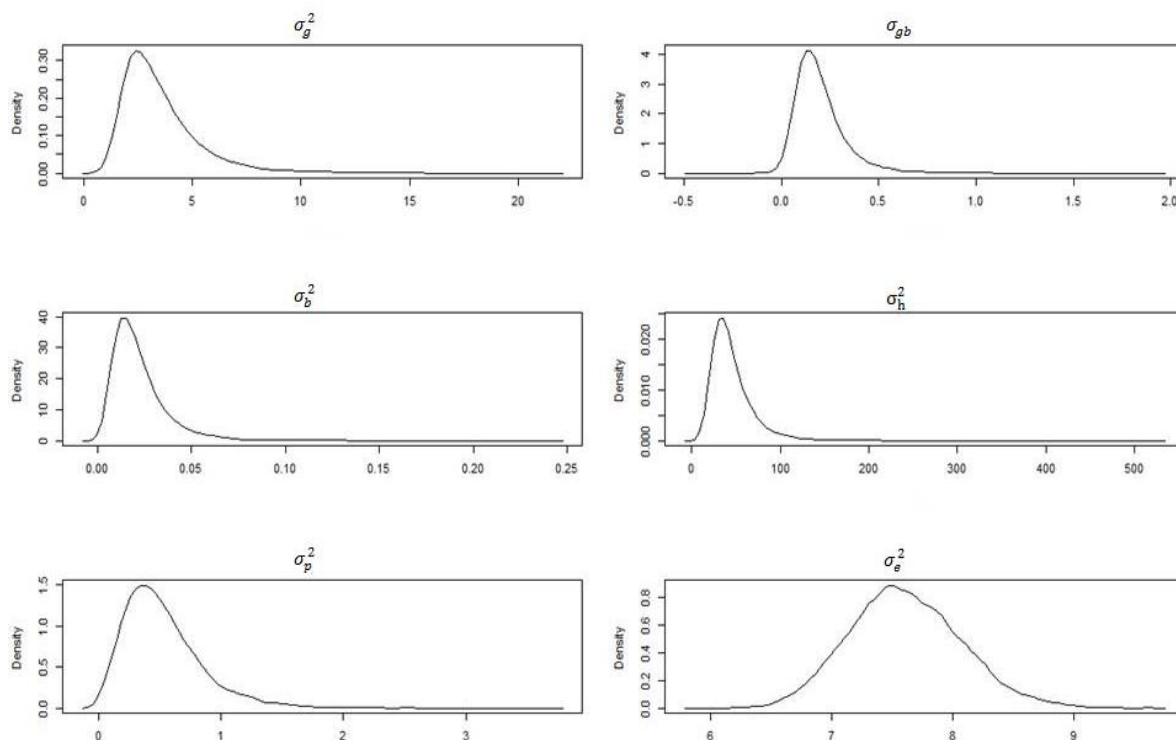


Figura 3. Gráfico da função da distribuição de probabilidade para os componentes de variância e co-variância para o **primeiro corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

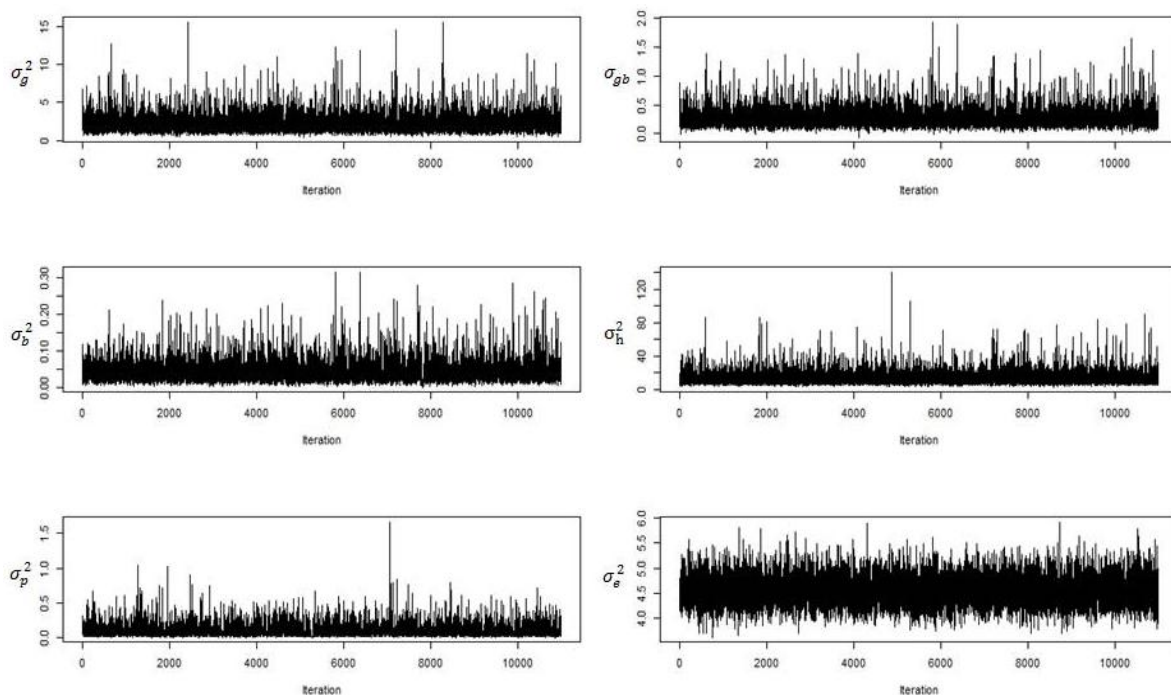


Figura 4. Gráfico de traço dos componentes de variância e co-variância estimados vs. iterações de cadeia de Markov para o **segundo corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

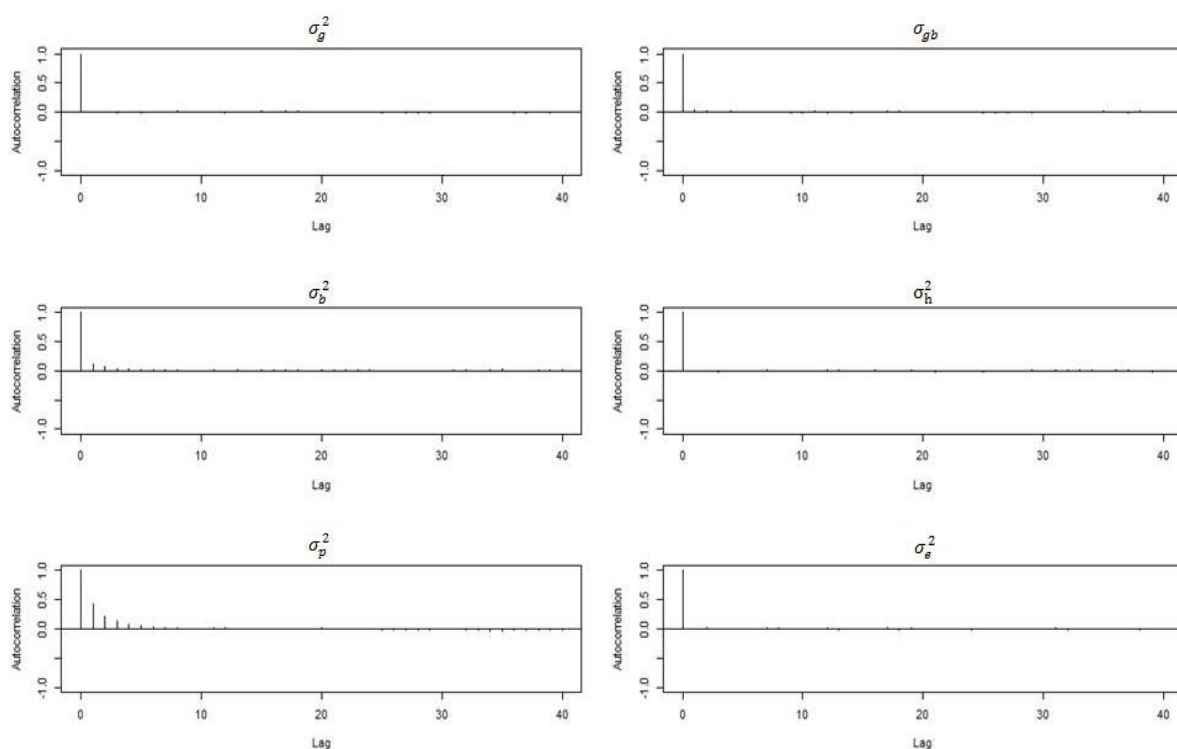


Figura 5. Gráfico das autocorrelações dos componentes de variância e co-variância estimados para o **segundo corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

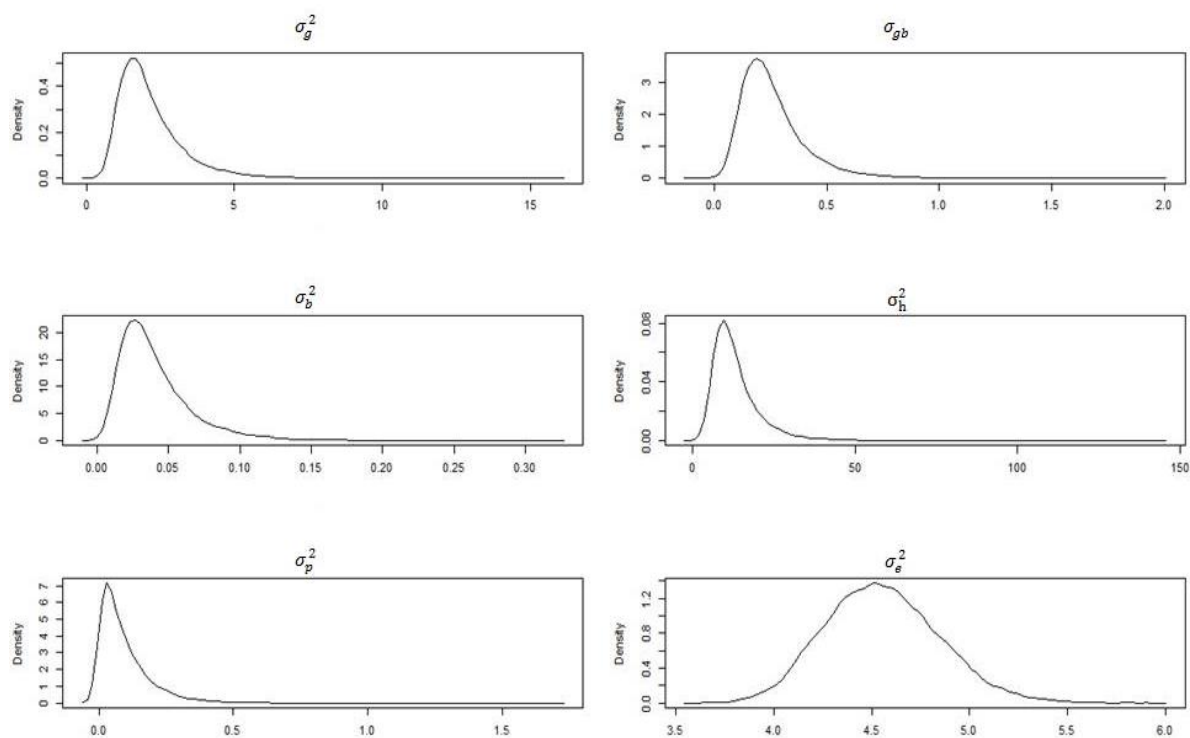


Figura 6. Gráfico da função da distribuição de probabilidade para os componentes de variância e co-variância para o **segundo corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão

genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

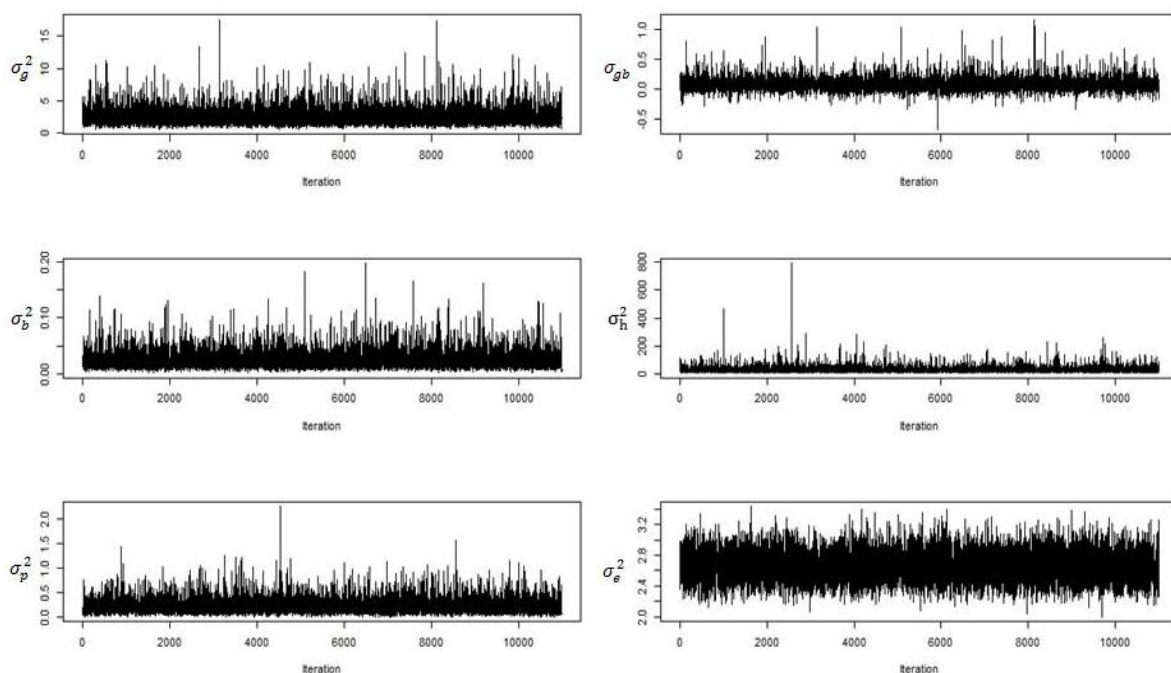


Figura 7. Gráfico de traço dos componentes de variância e co-variância estimados vs. iterações de cadeia de Markov para o **terceiro corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

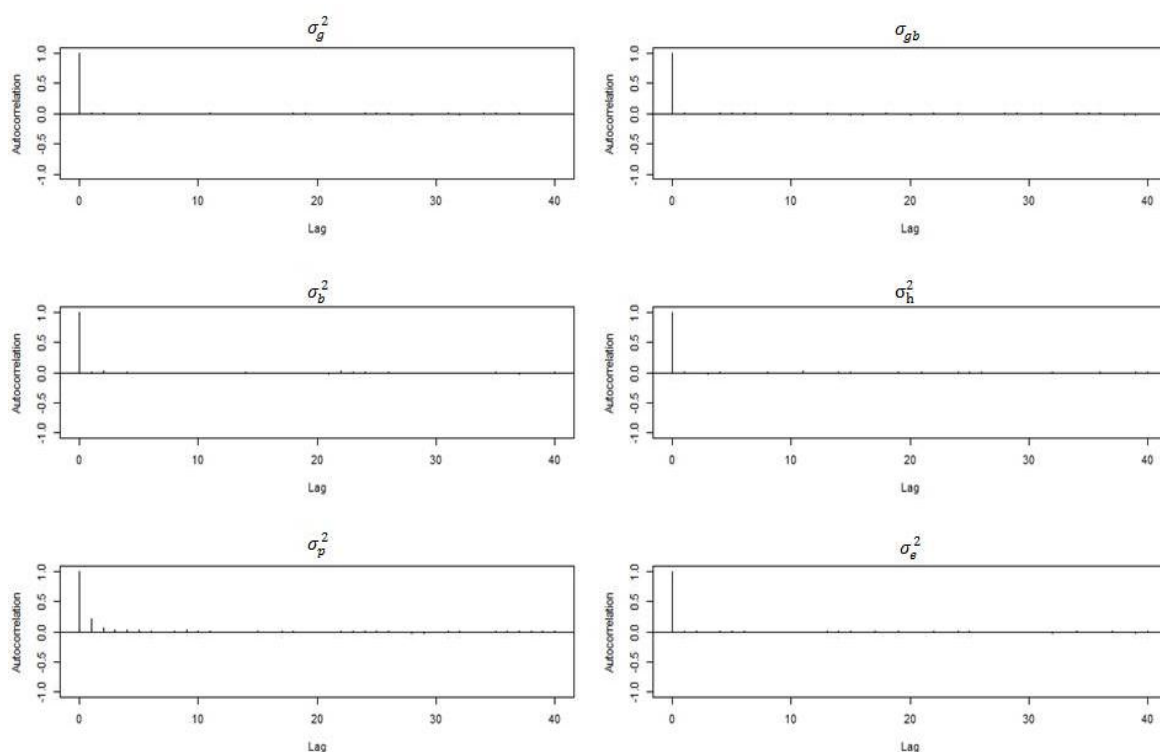


Figura 8. Gráfico das autocorrelações dos componentes de variância e co-variância

estimados para o **terceiro corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

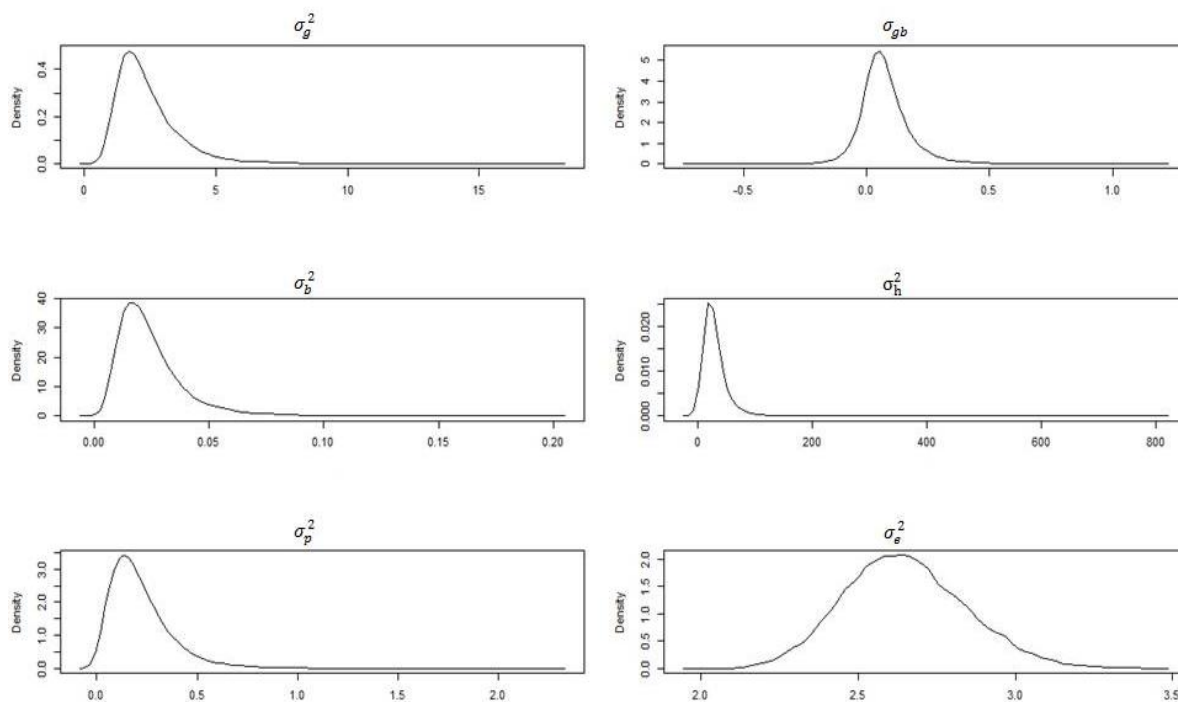


Figura 9. Gráfico da função da distribuição de probabilidade para os componentes de variância e co-variância para o **terceiro corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

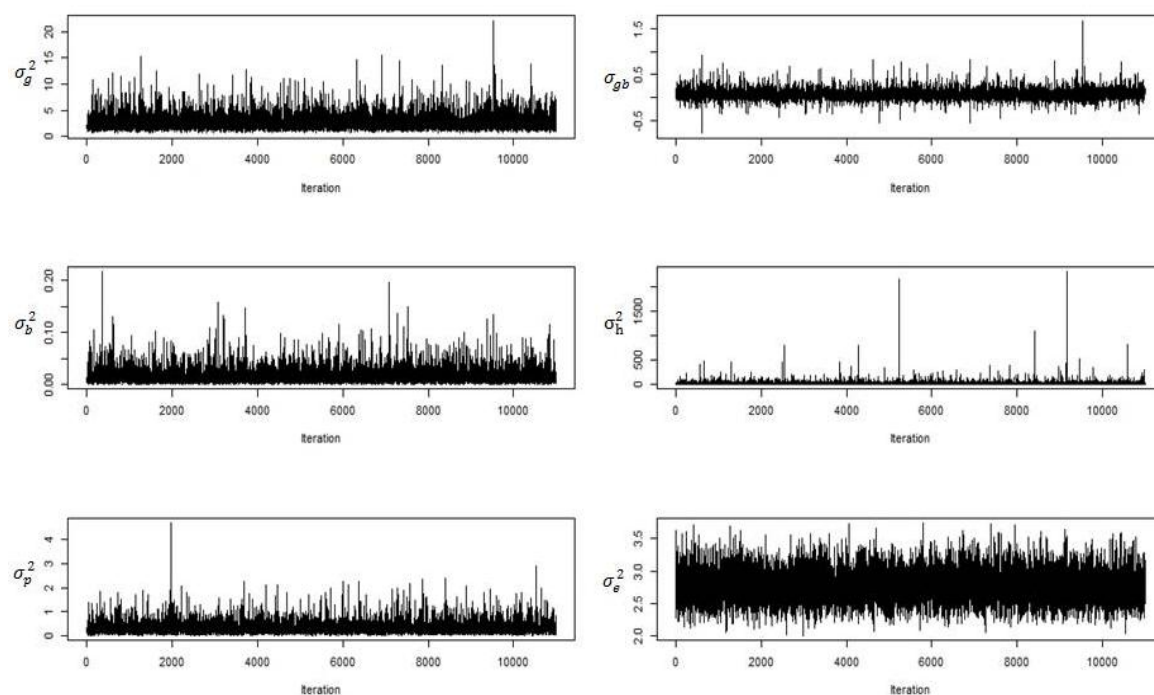


Figura 10. Gráfico de traço dos componentes de variância e co-variância estimados vs. iterações de cadeia de Markov para o **quarto corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do

coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

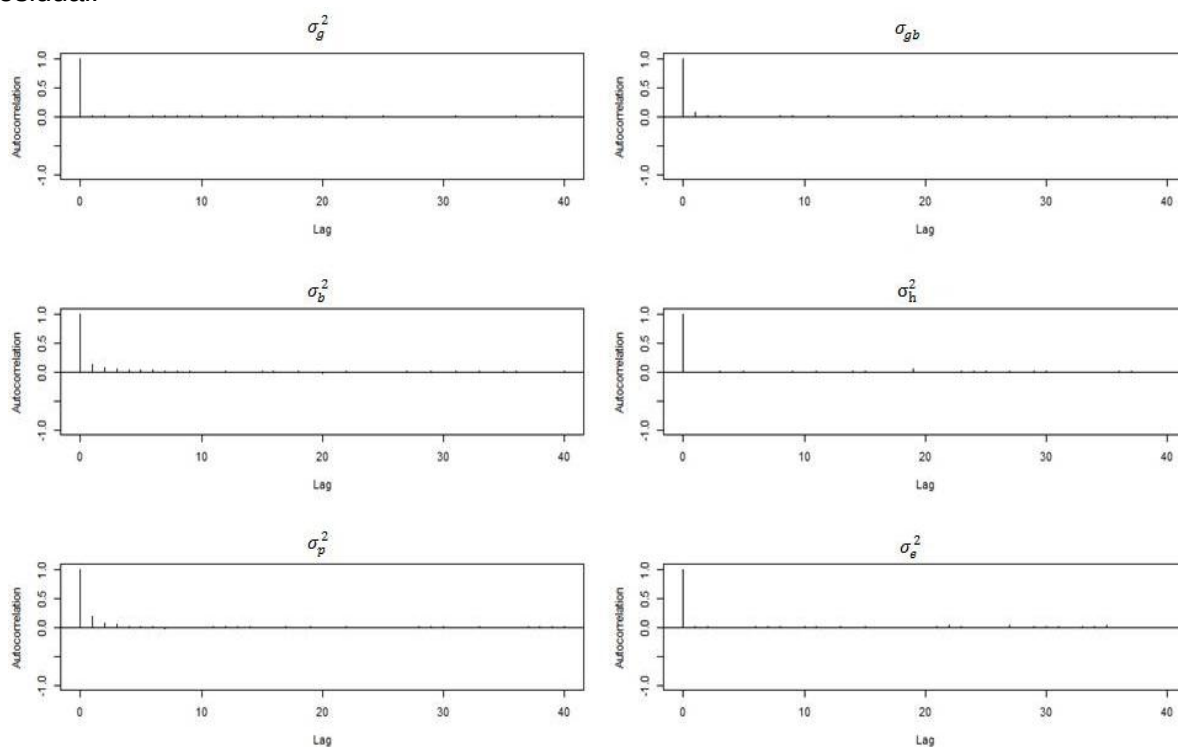


Figura 11. Gráfico das autocorrelações dos componentes de variância e co-variância estimados para o **quarto corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

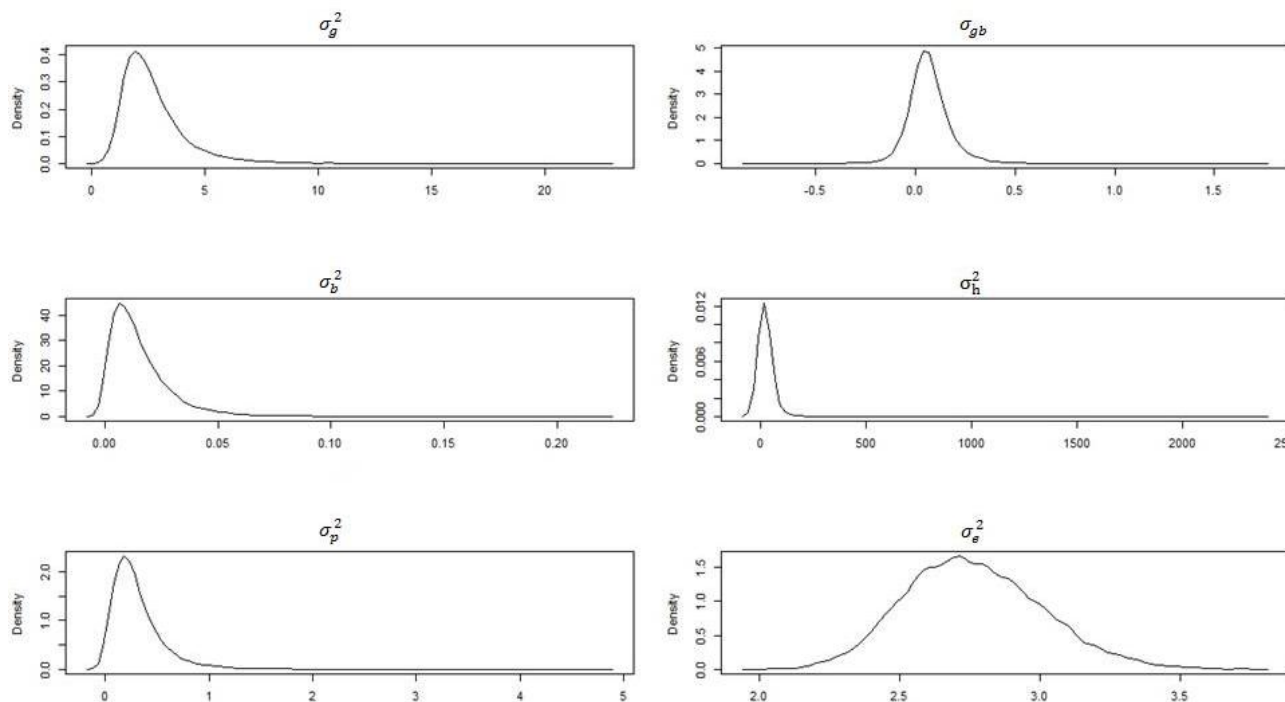


Figura 12. Gráfico da função da distribuição de probabilidade para os componentes de variância e co-variância para o **quarto corte**. σ_g^2 : Variância genotípica, σ_b^2 : variância do coeficiente de regressão genotípico, σ_{gb} : covariância genotípica e do coeficiente de regressão genotípico, σ_h^2 : variância do efeito do ambiente, σ_p^2 : variância do efeito de bloco, σ_e^2 : variância do efeito residual.

Tabela 13. Diagnóstico de Geweke para as variâncias e co-variância nos quatro cortes analisados.

Corte		σ_g^2	σ_{gb}	σ_b^2	σ_h^2	σ_p^2	σ_e^2
1	Z-Score	-0.7545	-1.2835	-1.4376	1.1842	-0.0023	-0.6550
	p-value	0.4506	0.1993	0.1506	0.2363	0.9982	0.5125
2	Z-Score	-0.3907	-1.4957	-1.1361	0.0190	-0.2927	-0.7384
	p-value	0.6960	0.1347	0.2559	0.9848	0.7697	0.4602
3	Z-Score	0.3402	0.1862	-0.7200	0.4270	-1.0729	1.7909
	p-value	0.7337	0.8523	0.4715	0.6693	0.2833	0.0733
4	Z-Score	1.1115	0.7999	1.2468	-0.4134	0.2966	-0.7111
	p-value	0.2663	0.4238	0.2125	0.6793	0.7668	0.4770

Variância genotípica (σ_g^2), Covariância genotípica e coeficiente de regressão da norma de reação (σ_{gb}), Variância do coeficiente de regressão genotípico (σ_b^2), Variância do efeito do ambiente (σ_h^2), Variância do efeito de bloco (σ_p^2) e Variância do efeito residual (σ_e^2).