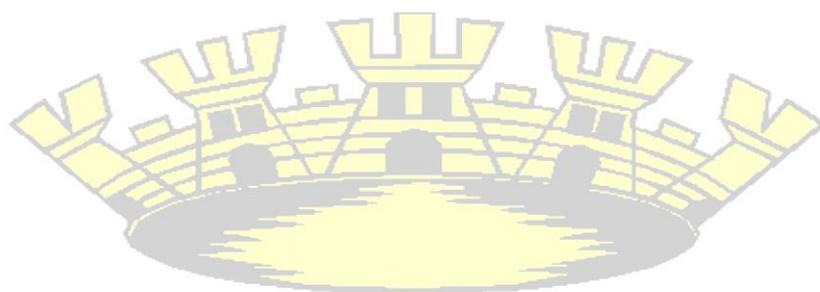


CLÉIA FLORENTINO DOS SANTOS



**USO DE DIFERENTES METODOLOGIAS ESTATÍSTICAS NO  
MELHORAMENTO DO AMENDOIM FORRAGEIRO**



RIO BRANCO – AC

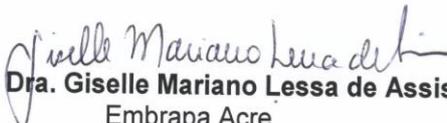
2014

CLÉIA FLORENTINO DOS SANTOS

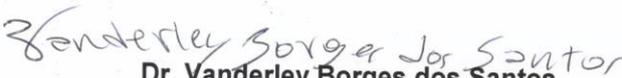
**USO DE DIFERENTES METODOLOGIAS ESTATÍSTICAS NO  
MELHORAMENTO DO AMENDOIM FORRAGEIRO**

Dissertação apresentada ao Programa de Pós-graduação em Agronomia, Área de Concentração em Produção Vegetal, da Universidade Federal do Acre, como parte das exigências para a obtenção do título de Mestre em Agronomia.

APROVADA em 14 de abril de 2014

  
**Profa. Dra. Giselle Mariano Lessa de Assis**  
Embrapa Acre  
Orientadora

  
**Dr. Jacson Rondinelli da Silva Negreiros**  
Embrapa Acre  
Membro

  
**Dr. Vanderley Borges dos Santos**  
Universidade Federal do Acre  
Membro

RIO BRANCO - AC  
2014

Ao meu filho, Rafael Viana  
dos Santos, por ser a luz da minha vida, à  
minha irmã Cristina Florentina dos Santos,  
por acreditar na minha capacidade de  
vencer, aos meus pais Francisco Augusto  
dos Santos e Joana Florentino dos Santos  
por ter me ensinado sempre os caminhos  
certos,  
Dedico

## AGRADECIMENTOS

Primeiramente a Deus, por ter me concedido a vida.

Aos meus orientadores, Giselle Mariano Lessa de Assis e José Marques Carneiro Júnior, pelos ensinamentos, paciência e compreensão.

Aos professores Jorge Ferreira Kusdra, Sebastião Elviro, Regina Felix, Vanderley Borges, Lidianne Assis, Marcos Deon e Leonardo Bering, pelo conhecimento adquirido e as palavras de incentivo.

À Universidade Federal do Acre, pela oportunidade de dar continuidade à minha formação.

À CAPES pelo apoio financeiro concedido.

À pesquisadora Dra. Patrícia Flores pelas dicas científicas e amizade.

Ao pesquisador Dr. Romeu de Carvalho A. Neto pelas palavras de incentivo, sempre bem vindas e pela amizade.

Aos meus irmãos Clemilse, Cláudio, Carlos, Clives e Magiri Maria (sobrinha e irmã), por estarem sempre presentes na minha vida.

Às minhas cunhadas Cirlândia Ribeiro, Macelina Gomes e aos cunhados Renato Xavier e Moacir Piveta.

À minha prima Mareilse de Albuquerque e ao José Ângelo pelo apoio nas horas de precisão.

Aos meus amigos Ana Ribeiro, José Marlon, Helen Sandra, Priscila Ferreira Wolter, Idianara Ribeiro, Alana, Hemithon Bandeira, Michele e Adilson, por me apoiarem sempre.

Aos colegas de curso Karina Galvão, Maisa Bravin, Gisley, Irene Ferro, Cassiano, Antônio Carlos, Romário Boldt, Noelene Thandara, João Paulo, Ana Cláudia, Maria Isabel, Isabela Santana, Márcia Capistrano pelo apoio e troca de conhecimentos, pelos momentos de descontração.

**Muito obrigada!**

## RESUMO

Este trabalho teve como objetivo verificar o efeito do uso de diferentes metodologias estatísticas de análise de dados na estimação de parâmetros genéticos e na seleção de amendoim forrageiro, utilizando dados simulados. A simulação dos valores fenotípicos baseou-se em dados experimentais de produção de matéria seca (Kg/ha) de amendoim forrageiro, obtidos a partir de avaliações agrônômicas realizada na Embrapa Acre. Utilizou-se o sistema computacional SAS<sup>®</sup> para simulação dos dados. Foram simuladas oito populações considerando-se quatro tamanhos distintos (4, 10, 30 e 100 genótipos), com dois níveis de desbalanceamento (0 e 20%). Foi considerada a realização de oito cortes, em delineamento de blocos ao acaso com seis repetições. Utilizou-se para análise dos dados: a metodologia de modelos mistos, em que os componentes de variância foram estimados pelo método da máxima verossimilhança restrita (REML) e os valores genotípicos foram preditos pelo método do melhor preditor linear não viesado (BLUP); e a metodologia de quadrados mínimos, por meio da análise de variância (ANOVA) em esquema de parcelas subdivididas. Para verificação da homogeneidade da estrutura da matriz de covariância residual entre cortes realizou-se teste de esfericidade de Mauchly. Foi calculada a correlação de Spearman e o Quadrado Médio do Erro para avaliar o ranqueamento e a acurácia da predição dos valores genéticos obtidos pela metodologia de modelos mistos. Os testes de esfericidade foram significativos a 5% de probabilidade pelo teste Qui-quadrado para todas as populações. As estimativas dos parâmetros genéticos obtidas pelas metodologias REML e ANOVA foram similares. O uso da metodologia REML/BLUP conduziu a maiores ganhos de seleção quando comparada à metodologia de quadrados mínimos, exceto para populações com quatro genótipos. Para populações maiores foram necessários um menor número de cortes para prever o valor genotípico dos indivíduos pelo método REML/BLUP com maior acurácia. Observou-se que o nível de desbalanceamento de 20% não interferiu na classificação dos genótipos superiores para as condições simuladas neste trabalho.

Palavras chave: *Arachis pintoi*. dados longitudinais. simulação computacional.

## ABSTRACT

This study aimed to verify the effect of using different statistical methods of data analysis in the estimation of genetic parameters and selection of forage peanut using simulated data. The simulation of phenotypic values was based on experimental data of dry matter yield (kg/ha) of forage peanut obtained from agronomic evaluations conducted at Embrapa Acre. The SAS<sup>®</sup> computer system was used for simulation data. Eight populations were simulated by considering four different sizes (4, 10, 30 and 100 genotypes), with two levels of unbalanced data (0 and 20%). Eight harvests were considered in a randomized block design with six replications. It was used for data analysis: the mixed model methodology, where the variance components were estimated by the restricted maximum likelihood method (REML) and the genotypic values were predicted by the best linear unbiased predictor method (BLUP); and the method of least squares, using the analysis of variance (ANOVA) in a split plot design. To check the homogeneity of the structure of the residual covariance matrix between harvests, the Mauchly's sphericity test was performed. Spearman Correlation and Mean Square Error were calculated to assess the ranking and accuracy of prediction of breeding values obtained by mixed model methodology. Tests of sphericity were significant at 5% probability by Chi-square test for all populations. Estimates of genetic parameters obtained by REML and ANOVA methods were similar. The use of the methodology REML/BLUP led to greater gains in selection when compared to the least squares method, except for populations with four genotypes. For larger populations fewer harvests were needed to predict the genotypic value of individuals by REML/BLUP method with higher accuracy. It was observed that the level of unbalance data of 20% did not interfere in the classification of superior genotypes for these simulated conditions in this work.

Keywords: *Arachis pintoii*. computer simulation. longitudinal data.

## LISTA DE FIGURAS

- Figura 1 - Etapas que compõem o Programa de Melhoramento Genético do Amendoim Forrageiro, desde a introdução de germoplasma até o lançamento da nova cultivar..... 19
- Figura 2 - Porcentagem de erro das estimativas dos componentes de variância e componentes quadráticos de genótipo, obtidos pelas metodologias REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas..... 48
- Figura 3 - Porcentagem de erro das estimativas dos componentes de variância e componentes quadráticos da interação genótipo x corte, obtidos pelas metodologias REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas..... 52
- Figura 4 - Porcentagem de erro das estimativas dos componentes de variância residual, obtidos pelas metodologias REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas..... 55
- Figura 5 - Porcentagem de erro das estimativas de herdabilidade e coeficiente de determinação genotípico, obtidos pelas metodologias REML e ANOVA, em populações com diferentes tamanhos, balanceadas e desbalanceadas..... 57

## LISTA DE TABELAS

Tabela 1 -	Populações simuladas pelo método Monte Carlo, considerando diferentes tamanhos e níveis de desbalanceamento.....	35
Tabela 2 -	Desvios padrão simulados para os diferentes tamanhos de populações.....	36
Tabela 3 -	Teste de esfericidade de Mauchly (W) para diferentes tamanhos de populações balanceadas, com base no teste de Qui-quadrado ( $\chi^2$ ).....	43
Tabela 4 -	Teste de esfericidade de Mauchly (W) para diferentes tamanhos de populações desbalanceadas, com base no teste de Qui-quadrado ( $\chi^2$ ).....	43
Tabela 5 -	Média e desvio padrão (DP) das estimativas de parâmetros genéticos obtidas pelos métodos REML e ANOVA para diferentes tamanhos de populações balanceadas.....	46
Tabela 6 -	Média e desvio padrão (DP) das estimativas de parâmetros genéticos obtidas pelos métodos REML e ANOVA para diferentes tamanhos de populações desbalanceadas.....	47
Tabela 7 -	Média da Correlação de Spearman (CS) e Quadrado Médio do Erro (QME) entre valores genotípicos reais e preditos pela metodologia REML/BLUP de populações com diferentes tamanhos, balanceadas e desbalanceadas.....	59

Tabela 8 -	Média e desvio padrão (DP) de parâmetros genéticos estimados pela metodologia REML/BLUP de populações balanceadas em função do número de cortes.....	62
Tabela 9 -	Correlação de Spearman (CS) e Quadrado Médio do Erro (QME) entre valores genotípicos reais e preditos pela metodologia REML/BLUP de diferentes populações balanceadas em função do número de cortes.....	65
Tabela 10 -	Média ( $\bar{X}$ ) e desvio padrão (DP) do ganho genético real da produção de matéria seca (Kg/ha) com a seleção em populações de diferentes tamanhos e dois níveis de desbalanceamento, utilizando o teste de Tukey .....	67
Tabela 11 -	Média ( $\bar{X}$ ) e desvio padrão (DP) de ganho genético real da produção de matéria seca (Kg/ha) com a seleção em populações de diferentes tamanhos e dois níveis de desbalanceamento, utilizando a metodologia REML/BLUP.....	67

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>12</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA.....</b>	<b>14</b>
2.1	AMENDOIM FORRAGEIRO.....	14
2.1.1	Melhoramento do amendoim forrageiro.....	16
2.2	EFEITOS FIXO E ALEATÓRIO NO MODELO ESTATÍSTICO NO MELHORAMENTO VEGETAL.....	20
2.3	MÉTODO DE QUADRADOS MÍNIMOS.....	21
2.4	METODOLOGIA DE MODELOS MISTOS.....	23
2.5	ANÁLISE DE DADOS COM MEDIDAS REPETIDAS.....	25
2.6	ANÁLISE DE DADOS DESBALANCEADOS.....	29
2.7	NÚMERO DE CORTES OU MEDIÇÕES NO MELHORAMENTO DE PLANTAS FORRAGEIRAS.....	30
2.8	SIMULAÇÃO DE DADOS.....	31
2.8.1	Uso da simulação de dados no melhoramento genético.....	33
<b>3</b>	<b>MATERIAL E MÉTODOS.....</b>	<b>35</b>
3.1	SIMULAÇÃO DOS DADOS.....	35
3.2	ANÁLISE DOS DADOS.....	37
3.2.1	Teste de esfericidade.....	37
3.2.2	Método de Quadrados Mínimos (LS).....	37
3.2.3	Metodologia de Modelos Mistos.....	38
3.2.4	Avaliação genética sob diferentes números de cortes.....	40
3.3	CRITÉRIOS DE AVALIAÇÃO .....	40
3.3.1	Acurácia das estimativas de parâmetros genéticos.....	41
3.3.2	Correlação entre valores genotípicos.....	41
3.3.4	Quadrado Médio do Erro (QME).....	42
3.3.5	Acurácia da predição dos valores genotípicos em função do número de cortes.....	42

3.3.6	Ranqueamento de genótipos e ganho de seleção.....	42
<b>4</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>43</b>
4.1	TESTE DE ESFERICIDADE.....	43
4.2	ESTIMATIVA DE PARÂMETROS GENÉTICOS.....	45
4.2.1	Varição do genótipo.....	45
4.2.2	Varição da interação genótipo x bloco.....	50
4.2.3	Varição da interação genótipo x corte.....	51
4.2.4	Varição residual.....	54
4.2.5	Herdabilidade e coeficiente de determinação genotípico.....	56
4.3	CORRELAÇÃO DE SPEARMAN E QUADRADO MÉDIO DO ERRO.....	59
4.4	AVALIAÇÃO GENÉTICA SOB DIFERENTES NÚMEROS DE CORTES.....	60
4.4.1	Estimativa de parâmetros genéticos.....	60
4.4.2	Correlação de Spearman e quadrado médio do erro.....	64
4.5	SELEÇÃO COM BASE EM QUADRADOS MÍNIMOS E REML/BLUP.....	66
<b>5</b>	<b>CONCLUSÕES.....</b>	<b>72</b>
	<b>REFERÊNCIAS.....</b>	<b>73</b>

## 1 INTRODUÇÃO

O melhoramento genético de plantas tem como objetivo a obtenção de novas cultivares mais produtivas e melhor adaptadas ao meio ambiente, que possuam características que solucionem problemas ou corrijam deficiências específicas de cultivares existentes.

Em relação ao melhoramento de plantas forrageiras, o processo de obtenção de cultivares superiores é bastante complexo, considerando que, nessas espécies, busca-se não apenas indivíduos mais produtivos e adaptados ao meio, mas também que proporcionem maior eficiência na produção de proteína e produtos animais de alto valor agregado.

As plantas forrageiras tropicais mais utilizadas no Brasil são perenes, de forma que a identificação dos melhores indivíduos requer mais de uma avaliação na mesma planta ou na mesma parcela. Esse processo gera dados longitudinais no tempo. O amendoim forrageiro (*Arachis pintoii* Krapov. & W.C. Greg.), leguminosa forrageira utilizada em pastagens consorciadas também é uma planta herbácea perene, cuja avaliação agronômica exige a obtenção de dados na mesma unidade experimental ao longo do tempo.

A análise de dados longitudinais é caracterizada, principalmente, pela estrutura das correlações existentes entre as medições e a presença ou não de heterogeneidade de variâncias. Estes aspectos levam à complexidade dos modelos estatísticos para análise de dados com estas características e tendem a provocar distorções na seleção dos indivíduos superiores, quando os modelos não condizem com a natureza biológica do caráter avaliado, ou ainda, se utilizam métodos pouco robustos para estimação de parâmetros.

Dentre os métodos utilizados para avaliação de medidas repetidas destacam-se o método de quadrados mínimos (LS), com esquema em parcelas subdivididas no tempo, utilizando-se a análise de variância (ANOVA); a análise multivariada, em que cada medida repetida no tempo é considerada uma variável; e a metodologia de modelos mistos, em que os componentes de variância são usualmente estimados pelo método da Máxima Verossimilhança Restrita (REML) e os valores genéticos são preditos pelo Melhor Preditor Linear não Viesado (BLUP).

O método dos mínimos quadrados, por meio da ANOVA, tradicionalmente, é a metodologia mais utilizada na avaliação de forrageiras

tropicais; no entanto, faz suposições mais restritivas quanto à estrutura de covariância dos dados, podendo afetar as inferências em diversas situações. Adicionalmente, pode gerar estimativas de componentes de variância negativas, além de suas propriedades serem pouco conhecidas em condições de desbalanceamento de dados. A técnica multivariada é uma boa opção, porém, o excesso de parâmetros a serem estimados dificulta a análise dos dados. O emprego da metodologia de modelos mistos no melhoramento vegetal vem crescendo consideravelmente nos últimos 15 anos, principalmente devido ao aumento da capacidade de processamento computacional, desenvolvimento de softwares amigáveis e divulgação da metodologia entre os melhoristas.

A escolha inadequada da metodologia pode levar a inferências imprecisas quanto à seleção dos melhores genótipos, ocasionando menores ganhos com a seleção e frustração do programa de melhoramento genético. Tendo em vista que os pesquisadores utilizam o método de quadrados mínimos com bastante frequência no melhoramento genético de plantas perenes, são necessários estudos que avaliem a eficiência dessa metodologia na seleção de genótipos superiores.

Para testar a eficiência de diferentes metodologias, a simulação de dados é bastante empregada, pois garante o conhecimento dos verdadeiros valores dos efeitos que compõem o fenótipo do indivíduo. A simulação busca criar um cenário mais próximo do que ocorre em um sistema real, sendo possível diagnosticar problemas ou fazer a escolha de metodologias que mais se adequam aos dados, com menor custo e maior rapidez. Assim, pesquisas que auxiliem na escolha do método mais adequado para seleção de genótipos superiores de plantas forrageiras tropicais são necessárias, de acordo com as particularidades de cada experimento.

Conforme o contexto apresentado, este trabalho teve como objetivo verificar o efeito do uso de diferentes metodologias estatísticas na estimação de parâmetros genéticos e na seleção de genótipos de amendoim forrageiro, utilizando-se a simulação computacional.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 AMENDOIM FORRAGEIRO

As pastagens nativas ou cultivadas são de grande importância para economia do país. Nos últimos anos observou-se um aumento da eficiência da produção de carne, levando o Brasil a ser o segundo maior produtor de carne bovina do mundo (USDA, 2013). Segundo o IBGE (2012), somente na região norte do país, o efetivo de bovinos manteve-se crescente todos os anos, considerando os últimos cinco anos. Destes, uma parcela de mais de 90% é alimentada em sistemas baseados exclusivamente a pasto. No entanto, estes sistemas apresentam baixa produtividade, mas conferem grande vantagem competitiva, pois permite a obtenção de produtos com baixo custo e boa qualidade (BARCELOS et al., 2001; EUCLIDES et al., 2001).

A baixa produtividade apresentada pelos sistemas de produção a pasto pode ser minimizada pela inserção de leguminosas nas pastagens, que proporciona um aumento da disponibilidade de nutrientes aos animais, além de outras vantagens. As leguminosas constituem o segundo maior e mais importante grupo de plantas forrageiras e são conhecidas pela alta capacidade de produção e qualidade nutritiva, sendo que sua qualidade é mantida por períodos de tempo superiores ao das gramíneas (ARAUJO et al., 2008; PEREIRA et al., 2001).

Existe grande diversidade de leguminosas forrageiras no mundo, em pastagens naturais e em áreas de cultivo, pouco estudadas e que poderão contribuir de forma significativa para produção animal através da sua utilização mais intensa (PEREIRA et al., 2001).

Dentre as leguminosas forrageiras de clima tropical destacam-se o *Stylosanthes guianensis* (estilosantes), *Leucena leucocephala* (leucena), *Arachis pintoi* (amendoim forrageiro).

O gênero *Arachis* pertence à tribo Aeschynomeneae, subtribo Stylosanthinae, família Fabaceae, e apresenta como característica que o diferencia dos demais gêneros da família, a produção subterrânea de frutos originários de flores aéreas. Está distribuído naturalmente pelo Brasil, Bolívia, Paraguai, Argentina e Uruguai com cerca de 60 a 80 espécies distribuídas por esses países (VALLS; SIMPSON, 1994).

As espécies *Arachis pintoii* e *Arachis repens* conhecidas como amendoim forrageiro, possuem os acessos mais promissores para utilização nos países de clima tropical. Estas espécies exclusivas da flora brasileira são perenes e de hábito de crescimento estolonífero, produzindo raízes nos nós, que ficam em contato com o solo ou em locais que estejam com elevada umidade (ASSIS et al., 2009).

O amendoim forrageiro tem sido recomendado para uso na alimentação animal, para recuperação de áreas degradadas e como cobertura do solo para fins ornamentais e de conservação (MIRANDA, 2008). Esta espécie consorciada com gramíneas é de grande importância para produção pecuária, pois aumenta o teor de proteína do volumoso oferecido aos animais, diversifica o ecossistema, incorpora nitrogênio e promove uma maior cobertura do solo, protegendo-o contra a erosão (ASSIS; VALENTIM, 2009).

A espécie *A. pintoii* se destaca por apresentar elevada produção de matéria seca de boa qualidade (Valle, 2001) e manter associações estáveis com gramíneas agressivas sob pastejo intensivo por mais de dez anos (KERRIDGE, 1995). Essas vantagens ocorrem em virtude de características peculiares como: a proteção do ponto de crescimento, que permite a manutenção de uma área foliar residual mesmo quando submetido a pastejo contínuo e intenso; a permanência na pastagem sob condições de manejo adversas; a resistência ao pastejo, pois seus estolões enraízam fortemente; e a boa reserva de sementes enterradas (ASSIS et al., 2008; MACHADO et al., 2005).

O seu uso em consorcio com pastagens tem promovido o aumento do ganho de peso animal por área, o que é de grande importância nos sistemas pecuários. Perez e Pizarro (2005) em pastagens consorciadas com *A. pintoii*, obtiveram ganho de peso anual em bovinos castrados variando de 130 a 200 kg.animal<sup>-1</sup> e de 250 a 630 kg.ha<sup>-1</sup>.ano<sup>-1</sup>, dependendo do período do ano e da gramínea associada. Segundo Valentim et al. (2001) o amendoim forrageiro produz forragem de alta qualidade, o que resulta em elevado consumo pelos animais em pastejo.

Lascano (1994) salienta que o valor nutritivo de *Arachis pintoii* é maior que os valores apresentados pela maioria das leguminosas tropicais de importância comercial. Valentim et al. (2003) observaram teores médios de proteína bruta variando de 17,9 a 21,7% em acessos de amendoim forrageiro,

em Rio Branco (AC). Santos (2012) encontrou teores médios de proteína bruta nas cultivares Belmonte e Mandobi de 22,77 e 22,38% respectivamente na época de transição águas/seca e 25,16 e 23,22% na época seca para as mesmas cultivares mencionadas anteriormente em Rio Branco (AC).

Miranda et al. (2003) afirmam que a contribuição maior do *Arachis* ao sistema de produção com pastagens consorciadas, é seu potencial em fixar nitrogênio atmosférico, quando em associação com bactérias fixadoras, resultando em forragem de melhor valor nutritivo, além da melhoria da fertilidade do solo. Os mesmos autores verificaram em acessos de *Arachis pintoi* e *Arachis repens* que os acessos BRA 31801 e BRA 31828 fixaram 87 e 90% respectivamente, de suas demandas nutricionais em nitrogênio. Outros autores como Thomas et al. (1997), em estudo com *Arachis pintoi* CIAT 17434 em dois solos da Colômbia também observaram que esta espécie pode fixar até 80% de suas demandas nutricionais em nitrogênio.

Apesar das vantagens proporcionadas pelo consórcio entre gramíneas e amendoim forrageiro relatadas por diversos autores, essa prática ainda é pouco utilizada pelos produtores. Dentre as causas que limitam o uso dessa técnica nos sistemas de produção pecuários estão: a dificuldade de convencimento dos produtores sobre os benefícios do seu uso, a falta de interação entre os obtentores das tecnologias com a iniciativa privada, para fins de multiplicação e distribuição de sementes ou propágulos, baixa produção de sementes, e conseqüente custo elevado de comercialização (VALENTIM et al., 2009).

No Estado do Acre, porém, a ampla divulgação e distribuição de mudas do amendoim forrageiro cv. Belmonte (VALENTIM et al., 2001) vêm ampliando a área de pastagens consorciadas na região. Atualmente, há 136.700 ha (EMBRAPA, 2013) de pastagens consorciadas no Estado, garantindo maior persistência desses pastos e aumento da qualidade da forragem ofertada aos animais.

### 2.1.1 Melhoramento do amendoim forrageiro

No Brasil, a alimentação dos rebanhos bovinos depende quase que exclusivamente de forrageiras tropicais, as quais são as principais fontes de energia e nutrientes para os animais (FIGUEIREDO, 2011).

Apesar da grande importância que essas forrageiras apresentam 50 a 70% das áreas de pastagens no Brasil estão em processo de degradação (DIAS FILHO, 2013), sendo esta, uma das causas da baixa produtividade dos rebanhos. Outro fator preocupante para cadeia produtiva é a perda da capacidade de produção das pastagens. Esta perda pode ser atribuída ao estabelecimento de cultivares não adaptadas para o local, ao manejo inadequado das pastagens, insuficiência ou inexistência de correção e manutenção da fertilidade do solo e à formação de pastagens com gramíneas puras (BARCELLOS et al., 2001; MACEDO, 2005; VALENTIM et al., 2001; VILELA et al., 2004;).

O aumento da produtividade dos rebanhos exige pastagens melhoradas, o que se traduz não apenas em cultivares produtivas e bem adaptadas, mas em maior diversidade para que se possa atender os requisitos de animais selecionados, para maior desempenho em todas as fases de desenvolvimento (VALLE et al., 2003).

No melhoramento de plantas forrageiras o objetivo não é somente conseguir uma planta altamente produtiva, mas obter uma maior eficiência na transformação da mesma em produto animal (PEREIRA et al., 2001). Esta eficiência na transformação da planta em carne, leite, couro ou lã, depende do consumo de alimento para sua nutrição, que em caso de pastagens consorciadas, está diretamente relacionado com a produtividade das pastagens, do seu valor nutritivo e persistência (ASSIS, 2010).

Estas particularidades das plantas forrageiras torna o melhoramento genético dessas culturas mais complexo e de longo prazo, levando cerca de seis a oito anos para o lançamento de uma nova cultivar. O processo de avaliação geralmente é realizado em três fases, sendo que na primeira fase é avaliado um grande número de genótipos, quanto a caracteres agrônômicos e nutricionais, por 2 a 3 anos. Em seguida avalia-se o efeito do animal sobre o pasto por 2 anos, e na terceira fase com número reduzido de genótipos, avalia-se o efeito da forrageira sobre o animal, como ganho de peso e produção de leite por animal por área, durante 2 ou 3 anos (ASSIS, 2010).

Segundo Simom et al. (2010) Unidades Descentralizadas da Embrapa conservam uma ampla variedade de leguminosas forrageiras em bancos de germoplasma, com mais de 6.000 acessos distribuídos em 47 gêneros, sendo que o gênero *Arachis* representa 9% desse total.

No Estado do Acre, está localizado o Banco Ativo de Germoplasma do Amendoim Forrageiro e o Programa de Melhoramento Genético do Amendoim Forrageiro coordenado pela Embrapa. Este programa foi iniciado em 2005, e tem como objetivo desenvolver cultivares propagadas por sementes e cultivares propagadas por mudas, com características agronômicas e bromatológicas de interesse, tais como: cobertura rápida do solo, elevada produção de matéria seca, alto valor nutritivo, resistência ao pisoteio e elevada persistência em sistemas consorciados (ASSIS, 2010).

Recentemente, foram estabelecidos descritores morfológicos para *Arachis pintoi* e definidos os intervalos para classificação das cultivares a partir da caracterização de acessos do Banco Ativo de Germoplasma, proporcionando a condução de ensaios de distinguibilidade, homogeneidade e estabilidade (DHE) desta espécie (ASSIS, 2010). No Banco Ativo de Germoplasma da Embrapa Acre há 139 genótipos de amendoim forrageiro, sendo empregadas diversas características na caracterização e avaliação desses acessos (ASSIS et al., 2012). As etapas que compõem o programa de melhoramento do amendoim forrageiro encontram-se representadas na Figura 1.

Entre os estudos com amendoim forrageiro, visando o melhoramento genético da espécie estão os de diversidade genética realizados por Ferreira et al. (2012), onde os autores avaliaram dez acessos de *Arachis spp.* e seis acessos de *A. pintoi* quanto à divergência nutricional. Este estudo mostrou que existe grande variabilidade na composição química desses acessos podendo ser selecionadas cultivares com maior qualidade nutricional. Menezes et al. (2012) avaliaram 12 genótipos de *A. pintoi*, quatro genótipos de *A. repens* e dois híbridos de *A. pintoi* x *A. repens* quanto à diversidade genética de características agronômicas e de composição química, mostrando que há variabilidade para características agronômicas de interesse, como produção de matéria seca. Estes estudos indicam alto ganho com a seleção de genótipos superiores desta espécie.

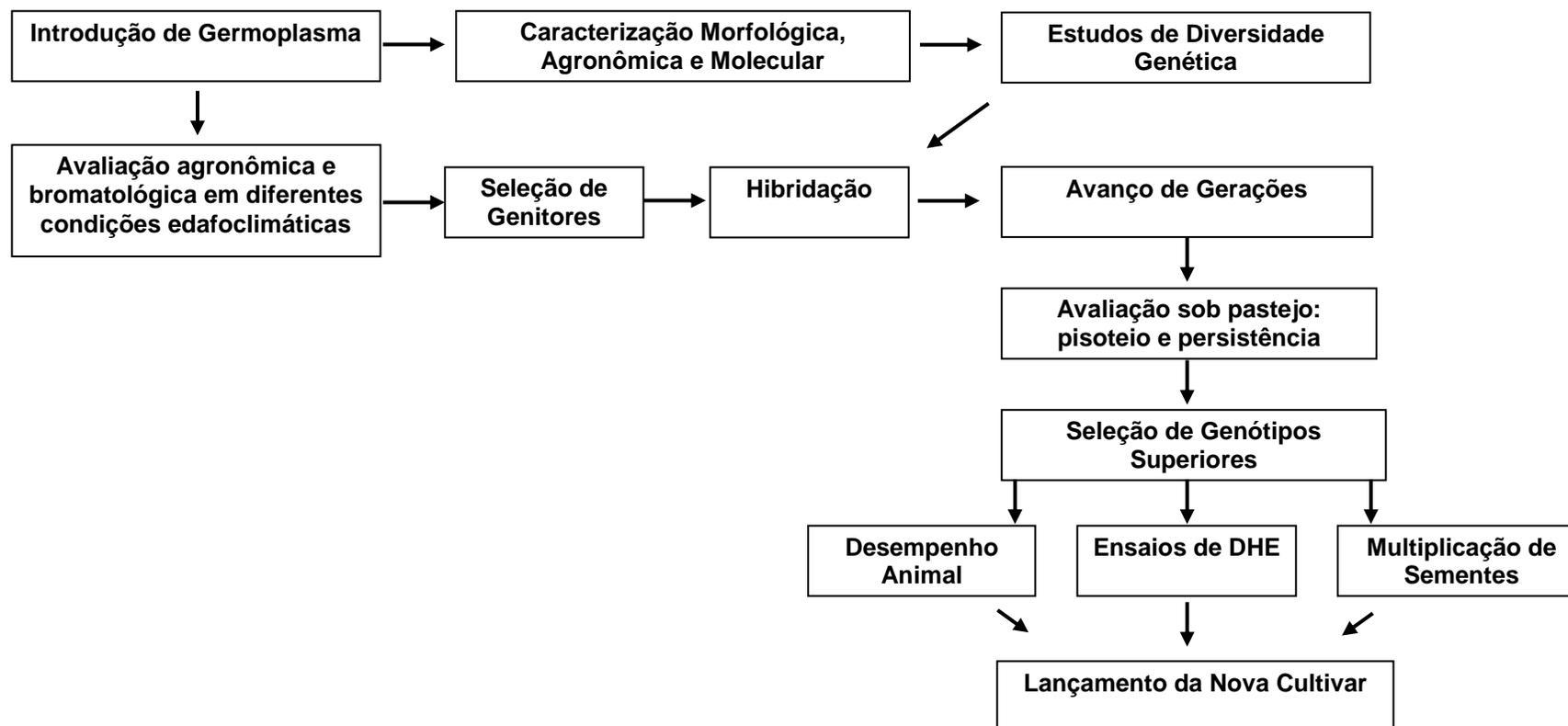


Figura 1- Etapas que compõem o Programa de Melhoramento Genético do Amendoim Forrageiro, desde a introdução de germoplasma até o lançamento da nova cultivar

Fonte: Assis (2010)

Adicionalmente, foram realizados estudos visando selecionar genótipos de amendoim forrageiro para cobertura do solo e produção de biomassa aérea no período de estabelecimento, utilizando-se mudas nas condições ambientais do Acre (ASSIS et al., 2008). Neste estudo, foram utilizados 21 genótipos, sendo 11 acessos de *A. pintoi*, quatro de *A. repens*, dois híbridos intraespecíficos de *A. pintoi*, um híbrido interespecífico de *A. pintoi* x *A. repens* e três cultivares de *A. pintoi*. Os autores identificaram genótipos superiores para cobertura do solo e produção de matéria seca simultaneamente.

Outro estudo de importância para maior sucesso na utilização do amendoim forrageiro em pastagens consorciadas foi realizado por Azevedo et al. (2014). Estes autores avaliaram o potencial de 110 genótipos quanto à associação com fungos micorrízicos arbusculares, pertencentes ao Banco Ativo de Germoplasma de Amendoim Forrageiro, localizado na Embrapa Acre. Os autores observaram que existe variabilidade genética entre os acessos estudados, quanto à riqueza e à promoção da esporulação de espécies de fungos micorrízicos arbusculares.

Apesar dos diversos estudos com amendoim forrageiro conduzidos por vários pesquisadores (Assis et al., 2008; Machado et al., 2005; Pereira et al., 2011;;; Valentim et al., 2003) e da importância inquestionável da implantação dessa leguminosa em pastagens, o modo de herança dos caracteres agrônômicos mais importantes é desconhecido e, devido à sua natureza quantitativa, estimativas de herdabilidade, repetibilidade e correlação dos caracteres devem ser obtidas (RESENDE et al., 2008).

Devem ser consideradas também as interações entre os efeitos do modelo, devido à realização de várias medições/cortes na mesma planta e a avaliação em vários locais. Essas condições, aliadas à perda de plantas ou de parcelas experimentais conduzem à complexidade na análise de dados do amendoim forrageiro.

## 2.2 EFEITOS FIXO E ALEATÓRIO DO MODELO ESTATÍSTICO NO MELHORAMENTO VEGETAL

O objetivo dos modelos estatísticos é procurar descrever a realidade de algum fenômeno natural. Em experimentos, por serem planejados, tem-se a resposta (fenótipos observados) em função de fatores previamente escolhidos

pelo pesquisador e de suas interações. Fatores principais no melhoramento de plantas referem-se aos genótipos, colheitas e ambientes. Outro aspecto importante na modelagem é a escolha da natureza dos efeitos do modelo, o qual pode ser fixo ou aleatório. No entanto, a escolha dos efeitos do modelo deve ser feita de acordo com os objetivos da pesquisa. Considerando o efeito como aleatório, a inferência é ampliada para população de referência, enquanto que para efeitos fixos, a inferência é restrita ao nível do fator em estudo (ANDRADE, 2013).

Para a escolha da metodologia de análise dos dados observados, é necessário que o pesquisador decida qual a natureza dos efeitos do modelo. Existem três tipos de modelagem com foco na seleção. A primeira denominada tipo I ou modelo fixo, que considera todos os fatores de efeito fixo, exceto o erro. O tipo II ou modelo aleatório, em que apenas a constante populacional é considerada fixa e o tipo III, ou modelagem mista, que inclui efeitos fixos e aleatórios, além da constante (fixa) e do erro (aleatório) (EISENHART, 1947; HENDERSON, 1975).

Considerar efeitos como fixos ou aleatórios pode alterar a estimação de componentes de variância e o ranqueamento dos genótipos (CAMARINHA FILHO, 2002; DUARTE; VENCONVSKY, 2001). Resende (2007) afirma que para fins de melhoramento genético, efeitos de genótipos devem ser considerados como aleatórios. O autor argumenta que a seleção nas diferentes etapas do melhoramento visa o desempenho futuro das cultivares, além do genótipo ser de natureza não observável. Na análise utilizando modelo misto é efetuada uma penalização (efeito *shrinkage*) da performance da média fenotípica do genótipo, ajustada para os demais efeitos ambientais controlados, presentes no experimento, penalizando as predições baseadas em pequenas amostras (RESENDE, 2002).

### 2.3 MÉTODO DE QUADRADOS MÍNIMOS

O procedimento de estimação de componentes de variância por meio da análise de variância (ANOVA) baseia-se na metodologia de quadrados mínimos. Este método permite trabalhar com certo nível de desbalanceamento e pressupõe amostragem aleatória e erro independente, que dificilmente se verifica em experimentos com plantas perenes (RESENDE, 2002).

Para estimativa dos componentes de variância, este método consiste em equacionar os quadrados médios da análise de variância com suas respectivas esperanças matemáticas. Para dados balanceados, esta metodologia possui propriedades altamente desejáveis (BARBIN, 1993), pois são não viesados e de variância mínima. Porém, Searle (1971) afirma que este método, na situação de dados desbalanceados, apresenta a propriedade de não vício, mas as pressuposições de normalidade dos dados, independência e homogeneidade dos erros não são asseguradas. Outra desvantagem é a possibilidade de obtenção de estimativas negativas de componentes de variância (RESENDE et al., 1996).

Caso haja desbalanceamento dos dados, a ANOVA conduz a estimativas imprecisas de componentes de variância e, conseqüentemente, à possibilidade de predições pouco acuradas dos valores genéticos (RESENDE et al., 2008).

Outro fato é que os testes que utilizam a metodologia de quadrados mínimos foram desenvolvidos sob a suposição de efeito de tratamento como fixo (STEEL; TORRIE, 1980), sendo, a rigor, inadequado o uso de tratamento aleatório em um modelo utilizando este método. Segundo Resende e Duarte (2007), os valores genotípicos das cultivares são variáveis aleatórias não observáveis, sendo essencial que o modelo estatístico utilizado para predição de valores genéticos considere o tratamento como efeito aleatório. Caso contrário, a seleção é fenotípica e não genética. Duarte (2000), Duarte e Venconvsky (2001), Resende (2002) e Smith et al. (2001) também enfatizam a necessidade de se considerar os materiais genéticos como efeitos aleatórios.

Bueno Filho (1997) encontrou diferenças no ordenamento de genótipos em ensaios florestais ao considerar o modelo como fixo ou aleatório no processo de predição de valores genéticos. Este resultado confirma a importância da definição do modelo como fixo ou aleatório, pois a escolha inadequada do modelo pode levar à seleção de indivíduos não superiores geneticamente.

O tamanho da amostra é outro aspecto a ser considerado quanto à escolha da metodologia de predição do valor genético. Em estudo realizado por Stein (1955), verificou-se que médias obtidas por quadrados mínimos só são adequadas quando se tem dois tratamentos em avaliação. Resende e Duarte

(2007) afirmam que os métodos tradicionais só são adequados para experimentos com até quatro tratamentos.

Vários são os estudos na área de melhoramento genético de plantas, realizados utilizando-se a metodologia de quadrados mínimos como: Léo et al. (2004), avaliando cultivares de alfafa; Cecon et al. (2008), em análise de medidas repetidas na avaliação de café *Conilon*; Daher et al. (2004), estimando parâmetros genéticos e coeficientes de repetibilidade em capim elefante; Cavalcante et al. (2012), estimando coeficientes de repetibilidade e parâmetros genéticos em genótipos de capim elefante; Souza Sobrinho et al. (2010), estimando coeficientes de repetibilidade e número ideal de cortes para seleção de *Urochloa ruziziensis*; Souza Sobrinho et al. (2010), avaliando a resistência de 94 clones de *B. ruziziensis* a duas espécies de cigarrinha das pastagens; Rocha et al. (2012), avaliando a eficiência da seleção para incremento do teor de óleo do pinhão manso.

## 2.4 METODOLOGIA DE MODELOS MISTOS

No melhoramento de plantas é comum o uso de análise estatística baseada em modelo fixo para estimação de parâmetros genéticos e comparação de médias de tratamentos. Em inúmeros casos, a modelagem mista é utilizada a rigor apenas na estimação de componentes de variância e para construção dos testes F apropriados na análise de variância (DUARTE; VENCOSKY, 2001).

Uma das análises no melhoramento de plantas, que tem sido utilizada mais recentemente, é a realizada por meio da metodologia de modelos mistos. Esta metodologia foi proposta inicialmente por Henderson (1949), e consiste na obtenção de valores genéticos considerados como sendo de efeitos aleatórios, corrigidos para os demais efeitos fixos contidos no modelo. Através dela pode-se obter o melhor preditor linear não viesado (BLUP, do inglês *Best Linear Unbiased Predictor*), que prediz o valor genético de cada indivíduo, além da obtenção do melhor estimador linear não viesado (BLUE, do inglês *Best Linear Unbiased Estimator*) dos efeitos fixos (HENDERSON, 1975)

A predição dos valores genéticos pelo BLUP requer o conhecimento prévio dos componentes de variância e covariância, que podem ser estimados por diferentes metodologias, dentre elas a máxima verossimilhança restrita

(REML, do inglês *Restricted Maximum Likelihood*) de Patterson e Thompson, derivados da função de verossimilhança (PATTERSON; THOMPSON, 1971).

O método REML/BLUP possibilita a análise de dados desbalanceados, com heterogeneidade de variâncias e permite o uso de efeitos aleatórios e fixos. Os efeitos de genótipo são sempre aleatórios, sendo que os demais efeitos do modelo podem ser fixos ou aleatórios.

No melhoramento de plantas, o método REML/BLUP passou a ser utilizado a partir da década de 1990. No Brasil, passou a ser usado em 1993 (RESENDE et al., 1996) e poucos trabalhos foram realizados antes de 2000 (BUENO FILHO, 1997). Dentre as razões que levam os melhoristas a não utilizarem predições baseadas em modelos mistos, estão a falta de vivência com estes métodos e a sua pequena divulgação (BUENO FILHO, 1997). Além disso, os efeitos prejudiciais da abordagem tradicional são tidos como mínimos a ponto de não recompensar os esforços com a adoção da metodologia de modelos mistos (ROSADO, 2003). Resende (2007a) destaca as principais vantagens do REML/BLUP, como: comparar indivíduos ou linhagens através do tempo e espaço; permitir a simultânea correção para os efeitos ambientais; realizar a estimação de componentes de variância e predição de valores genéticos; permitir estruturas complexas de dados (medidas repetidas, diferentes anos, locais e delineamentos); poder ser aplicado a dados balanceados e desbalanceados; permitir a utilização simultânea de um grande número de informações; permitir o ajuste de vários modelos alternativos, podendo-se escolher o que se ajusta melhor aos dados e ao mesmo tempo, ser parcimonioso (apresenta menor número de parâmetros).

Um modelo misto pode ser expresso matricialmente da seguinte forma:

$$y = X\beta + Zg + \varepsilon, \text{ em que:}$$

y: vetor de observações;

X: matriz de incidência de efeitos fixos;

Z: matriz de incidência dos efeitos aleatórios;

$\beta$ : vetor de efeitos fixos a serem estimados;

g: vetor de efeitos aleatórios a serem preditos;

$\varepsilon$ : vetor de erros aleatórios, associado a cada observação tal que  $\varepsilon \sim N(0, \sigma^2)$ .

As dimensões dessas matrizes e vetores dependem do delineamento experimental empregado.

Searle et al. (1992) afirmam que a preferência pelo BLUP se deve às suas propriedades estatísticas, que são superiores às propriedades dos estimadores pelo método dos quadrados mínimos.

A superioridade do BLUP em relação a outras metodologias tem sido também relatada em diversos estudos (RESENDE; DUARTE, 2007; RESENDE, 2008; SEARLE, 1992), sendo usado em trabalhos com plantas perenes (RESENDE, 2002; ROCHA et al., 2006a; 2006b; ROSADO et al., 2009) e também em culturas anuais (FRITSCH NETO, 2010; SANTOS, 2009).

Em trabalho realizado por Resende et al. (1996a), os autores concluem que o procedimento REML/BLUP, no caso de dados balanceados, tem o mesmo efeito que o tradicional método de quadrados mínimos. Porém, em dados desbalanceados e elevado número de genótipos, a diferença no ordenamento das médias dos genótipos pelos dois métodos é uma realidade (DUARTE; VENCOSKY, 2001).

Em trabalho realizado por Furtini (2011), os métodos REML/BLUP e quadrados mínimos levaram a resultados semelhantes. O autor comparou o método BLUP, seleção massal e massal estratificada na seleção de progênies de *Eucalyptus* e concluiu que as metodologias selecionam os clones de modo semelhante. Carvalho et al. (2008), comparando os métodos REML/BLUP e quadrados mínimos, na estimação de parâmetros e predição de valores genéticos de progênies de soja, constataram que os dois métodos não apresentam diferenças práticas.

Marcelino e lemma (2000) apresentaram estudo visando mostrar numericamente os principais métodos de estimação de componentes de variância. Utilizando modelos mistos com dados desbalanceados, os autores testaram os métodos I, II e III de Henderson, método da análise de variância (ANOVA), máxima verossimilhança (ML) e máxima verossimilhança restrita (REML) e concluíram que, em dados desbalanceados, somente o ML e REML, fornecem estimativas não negativas para os componentes de variância. Os autores ressaltam, porém, que a conclusão sobre o melhor método depende de estudos com simulação computacional.

## 2.5 ANÁLISE DE DADOS COM MEDIDAS REPETIDAS

Medidas repetidas são obtidas por repetidas mensurações de um mesmo caráter em um mesmo indivíduo ou na mesma parcela, e são frequentemente realizadas em plantas perenes, como fruteiras, árvores florestais, forrageiras, ou seja, indivíduos que tenham um ciclo de vida longo e precisam ser avaliadas diversas vezes no decorrer do tempo.

Em estudos dessa natureza, utilizando-se parcelas subdivididas no tempo, admite-se que as medidas tomadas nas diversas avaliações têm variâncias homogêneas e são igualmente correlacionadas. Nestes experimentos, os níveis do fator tempo não podem ser aleatorizados para seus intervalos. Dessa forma, a análise de variância usual pode não ser válida, pois com a falta de aleatorização, os erros correspondentes às respectivas unidades experimentais ou indivíduos podem ter uma matriz de covariância residual que não é igual à exigida para que a análise usual desse delineamento seja válida, ou seja, ocorrência de variâncias homogêneas (XAVIER, 2000). Fernandez (1991) também salienta o problema de que quando o experimento é sistematicamente arranjado, sem aleatorização, a análise de experimentos com medidas repetidas com delineamento em parcelas subdivididas, pode inflacionar a probabilidade de falsamente rejeitar a hipótese nula (erro tipo I).

Para que a análise de variância usual seja válida é necessário que a matriz de covariâncias entre as medidas repetidas seja homogênea e, em plantas perenes, isso raramente acontece (REZENDE, 2007). Para que a matriz de covariância seja homogênea, esta tem que ser do tipo simetria composta. A condição de simetria composta implica que a variável aleatória seja igualmente correlacionada e tenha variâncias iguais, considerando as diferentes ocasiões.

Mauchly (1940) propôs o teste de esfericidade, que verifica se a população multivariada apresenta variâncias iguais e correlações nulas. Caso a matriz apresente essa simetria será chamada de “esférica”. Caso não apresente simetria, a análise de variância por quadrados mínimos não é adequada, sendo necessária a escolha de outra forma de análise dos dados.

Uma condição mais geral da forma da matriz de covariância com variâncias iguais é descrita por Huynh e Feldt (1970). Essa condição denominada HUYNG-FELDT (H-F), é equivalente a especificar que as

variâncias da diferença entre pares de erros são todas iguais, e se as variâncias são todas iguais, então a condição é equivalente à de simetria composta (XAVIER, 2000).

Freitas et al. (2011), em trabalho realizado com 92 genótipos de alfafa e 20 cortes, testando os métodos de análise parcelas subdivididas no tempo e modelos mistos com PROC GLM e MIXED do programa *Statistical Analysis System* (SAS<sup>®</sup>) (SAS INSTITUTE, 2003) ressaltam que o modelo de análise em parcelas subdivididas só apresenta resultados corretos quando a condição de circularidade e esfericidade é atendida, ou seja, a matriz de covariância apresenta simetria composta (CS).

Uma alternativa a ser utilizada quando a matriz de covariância residual entre as várias medições não atende a condição de esfericidade, é a análise por meio de modelos mistos. Esta metodologia permite que a matriz de covariâncias seja modelada da forma que melhor represente os dados, levando em consideração se os dados são dependentes, independentes, correlacionados, ou ainda, apresentam outra relação que a matriz de covariância usual não consegue explicar (XAVIER, 2000).

Como exemplo aplicado em culturas perenes, pode-se citar o trabalho de Assis et al. (2009), que ao realizarem o teste de esfericidade em dados oriundos da avaliação de 35 genótipos de alfafa para a produção de matéria seca, verificaram que a condição de homogeneidade da matriz de covariâncias não foi atendida. Assim, os autores avaliaram cinco estruturas de matriz de covariância, sendo duas tradicionais (simetria composta e univariada) e três estruturas espaciais (SP(POW), SP(GAU) e SP(SPH)), por meio do SAS<sup>®</sup> utilizando o comando PROC MIXED. Para selecionar a matriz mais adequada, foram utilizados o Critério de Informação de Akaike (AIC) e Critério Bayesiano (SBC), sendo que os dois critérios utilizados indicaram que a melhor estrutura para análise dos dados foi a espacial SP(POW).

Freitas et al. (2011) salientam que existe mais de 40 tipos de matrizes de covariância residual implementadas no programa estatístico SAS<sup>®</sup>, que podem ser usadas de acordo com o conhecimento do pesquisador, quando a condição de esfericidade não é satisfeita. A seguir, são apresentadas algumas estruturas da matriz de covariância residual ( $\Sigma$ ) entre medições mais utilizadas (SAS INSTITUTE, 2003):

a. Componente de Variância:

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}, \text{ em que } \sigma^2 \text{ é a variância das diferentes medições.}$$

Esta estrutura impõe variâncias iguais nas  $n_i$  ocasiões e observações independentes. Envolve um único parâmetro.

b. Simetria Composta:

$$\Sigma = \begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 \end{bmatrix}, \text{ em que } \sigma^2 \text{ é a variância das diferentes medições e}$$

$\sigma_1$  é a covariância entre as medições.

Esta estrutura impõe variâncias iguais nas  $n_i$  ocasiões e mesma covariância entre medidas feitas em ocasiões distintas.

c. Simetria Composta Heterogêneas

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 \end{bmatrix}, \text{ em que } \sigma_i^2 \text{ representa o } i\text{-ésimo parâmetro}$$

da variância e  $\rho$  representa a correlação entre as medições, satisfazendo  $|\rho| < 1$ .

Impõe parâmetros de variâncias diferentes para cada elemento da diagonal principal, e utiliza a raiz quadrada desses parâmetros para compor os elementos fora da diagonal principal

d. Não estruturada:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 & \sigma_{13}^2 \\ \sigma_{21}^2 & \sigma_2^2 & \sigma_{23}^2 \\ \sigma_{31}^2 & \sigma_{32}^2 & \sigma_3^2 \end{bmatrix}, \text{ em que } \sigma_i^2 \text{ representa o } i\text{-ésimo parâmetro da}$$

variância.

Impõe variâncias distintas para cada uma das  $n_i$  ocasiões e covariâncias diferentes entre medidas feitas em ocasiões distintas.

e. Estrutura auto regressiva de primeira ordem:

$$\Sigma = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}, \text{ em que } \rho \text{ é a covariância entre medições.}$$

Impõe variâncias iguais nas diversas ocasiões e correlação decrescente entre o aumento do intervalo entre as ocasiões.

f. Estrutura auto regressiva de primeira ordem heterogênea:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 \end{bmatrix}, \text{ em que } \sigma_i^2 \text{ representa o } i\text{-ésimo}$$

parâmetro da variância.

É uma generalização da estrutura auto regressiva de primeira ordem, impondo variâncias e covariâncias diferentes.

Atualmente, alguns trabalhos têm sido realizados com o intuito de verificar a melhor estrutura de matriz de covariâncias para diferentes espécies vegetais, como: Mariguele et al. (2011), comparando formas de análise de medidas repetidas para o melhoramento genético da pinha (*Annona squamosa*), por meio de diferentes estruturas de matriz de covariância para os fatores progênie, parcela e resíduo; Cecon et al. (2008), trabalhando com a produção de 5 anos de 50 clones de café 'Conilon' avaliaram matrizes de covariância de simetria composta, auto-regressiva de primeira ordem e simetria composta heterogênea; Malheiros (1999), em estudo teórico, verificou que o modelo de matriz de covariância mais apropriado é o multivariado não estruturado.

## 2.6 ANÁLISE DE DADOS DESBALANCEADOS

A experimentação de campo está associada ao desbalanceamento de dados devido a variadas causas, como perda de plantas e parcelas, quantidades desiguais de sementes e mudas disponíveis, rede experimental com diferentes delineamentos experimentais e não avaliação de todas as combinações genótipo x ambiente.

A diminuição da taxa de sobrevivência em plantas forrageiras devido ao longo ciclo dessas culturas, por exemplo, gera dados desbalanceados para a

estimação dos parâmetros genéticos e predição dos valores genotípicos dos indivíduos. Há também o uso de delineamentos não ortogonais como BIB (blocos incompletos balanceados), PBIB (blocos incompletos parcialmente balanceados) e delineamentos aumentados (DUARTE, 2000), que são procedimentos planejados que preveem o não balanceamento dos dados.

Em dados desbalanceados, a análise de variância usual baseada na metodologia de quadrados mínimos para estimação de componentes de variância tem eficiência desconhecida (RESENDE et al. 1996). Duarte (2000) salienta que o desbalanceamento por si só, em qualquer delineamento, já é suficiente para não mais garantir a concordância na seleção por diferentes métodos, caso os tratamentos diferirem muito em número de repetições.

Segundo Rosado et al. (2009), o uso de diferentes metodologias para quantificar o ganho com a seleção de progênies provenientes de dados desbalanceados possibilitam classificações genotípicas diferenciadas, sendo que os ganhos preditos com a metodologia de modelos mistos é sempre superior. O mesmo autor relata que em dados balanceados e ortogonais a classificação dos genótipos não se altera.

Vários estudos na área de melhoramento vegetal têm indicado o método REML para estimativa de componentes de variância e a metodologia BLUP para predição de valores genéticos como a metodologia mais adequada para análise de dados desbalanceados, dentre os quais se pode citar: Faria (2008); Rosado et al. (2009); Sturion; Resende (2010).

## 2.7 NÚMERO DE CORTES OU MEDIÇÕES NO MELHORAMENTO DE PLANTAS FORRAGEIRAS

A avaliação de forrageiras tropicais requer várias medições e/ou cortes no mesmo indivíduo ao longo do tempo. Este tipo de avaliação tem por objetivo inferir sobre a capacidade do genótipo em repetir seu desempenho ao longo de sucessivas avaliações. Quanto maior o número de avaliações, maior gasto com mão de obra, materiais de consumo e tempo para o lançamento de novas cultivares.

O estudo do número de cortes necessários para inferir sobre o valor genético dos indivíduos em dados experimentais é feito por meio do coeficiente repetibilidade, que mede o grau de correlação entre as medidas repetidas no mesmo indivíduo no tempo ou no espaço (CRUZ et al., 2004).

O coeficiente de repetibilidade é amplamente usado pelos melhoristas de plantas como medida do limite superior da herdabilidade e da eficiência da predição do valor genético a partir de sucessivas medições no indivíduo (BASSO et al., 2009). Altos valores desse coeficiente, para quaisquer características, indicam que é possível predizer o valor real dos indivíduos com base em determinado número de medições (MASTUSCELLO et al., 2007).

O coeficiente de repetibilidade depende de estimativas fidedignas dos componentes de variância. A estimativa inadequada do número de cortes ou medições necessários pode levar a resultados de avaliações imprecisas nos programas de melhoramento genético de plantas forrageiras, além de maiores gastos de recursos quando estes forem recomendados acima do número necessário.

Atualmente, são inúmeros os estudos para avaliar o número de medições necessárias para estimar parâmetros genéticos com máxima confiança, e selecionar genótipos promissores de plantas forrageiras tropicais, dentre estes, estão: Cavalcante et al. (2012), avaliando genótipos de capim-elefante; Basso et al. (2009), avaliando acessos de *Brachiaria brizantha* Stapf; Daher et al. (2004), em estudo com capim-elefante; Souza Sobrinho et al. (2010), avaliando o número de cortes necessários para seleção de *Urochloa ruziziensis*.

Ganhos com a seleção são esperados em programas de melhoramento genético, porém os resultados das análises de dados são influenciados pela escolha do modelo de avaliação e pelas interações existentes entre os efeitos ambientais e de genótipos, causando resultados diferentes com a escolha de diferentes metodologias. Neste contexto, os estudos com simulação de dados levando em consideração todos esses aspectos são de grande importância para inferir com precisão o número de cortes necessários para predição do valor genético real em plantas forrageiras tropicais.

## 2.8 SIMULAÇÃO DE DADOS

A busca por eficiência nos programas de melhoramento genético é de fundamental importância, pois estes demandam alto custo e são de longa duração. Há várias alternativas que podem ser exploradas, contudo, a maioria demanda a realização de experimentos em variadas condições de acordo com a finalidade da pesquisa. Com as facilidades computacionais atuais, uma alternativa, de baixo custo e altamente confiável é a simulação de dados.

Esta ferramenta consiste em projetar um modelo de um sistema real e conduzir experimentos com esse modelo, no intuito de compreender o comportamento ou avaliar estratégias para operação do sistema utilizado. De forma prática, é um processo de criação de um modelo computadorizado tendo como objetivo a realização de experimentos numéricos, de tal forma que proporcione uma compreensão deste sistema submetido a um determinado conjunto de condições (KELTON et al., 2007).

Um aspecto que deve ser considerado é a eficiência do processo de simulação. Cuidados devem ser tomados para evitar erros devidos a problemas amostrais, escolhas inadequadas das distribuições de probabilidade nos eventos de natureza aleatória, simplificação inadequada da realidade e erros de implementação do sistema simulado (FERREIRA, 2001).

A qualidade e a validade de um modelo simulado são medidas pela proximidade entre os resultados obtidos pelo modelo e aqueles originados do sistema real. Assim, espera-se que a média das várias simulações de experimentos tenha o mínimo de variação possível, sendo condizentes ao que ocorre em sistemas reais (FREITAS FILHO, 2008).

Na área de melhoramento genético, esta técnica exige dos geneticistas o desenvolvimento de modelos biológicos que retratem da melhor maneira possível os fenômenos de interesse e, dos programadores, as rotinas para processamento adequado, impondo restrições para que a influência de certos fatores possa ser avaliada (CRUZ, 2001).

A simulação de dados é uma das áreas que mais se utiliza a geração de números aleatórios. Os primeiros indícios surgiram com a utilização do Método de Monte Carlo, por Von Neuman, em 1940, com blindagem de reatores nucleares (MORGAN, 1955). Esse método refere-se a processos de aproximação de valores esperados (integrais com respeito a uma distribuição de probabilidade) por meio de amostras (REZENDE, 2000).

De acordo com Ferreira (2001), no melhoramento de plantas, o uso da simulação é justificado quando: as soluções analíticas não existem ou o grau de dificuldade e o número de variáveis envolvidas não permitem a realização de inferências adequadas sobre o problema; pretende-se comparar a eficiência de um novo procedimento em relação a outros já existentes; os métodos que são rotineiramente empregados em algumas etapas de um programa de melhoramento propiciam qualquer aumento de sua eficiência. Também pode

ser aplicada ao estudo da herança por meio de componente de média, para avaliar a qualidade de estimadores e alertar os melhoristas quanto aos aspectos restritivos de uso dos estimadores (FERREIRA, 2008).

Em genética e melhoramento de plantas, a maioria dos problemas possui natureza aleatória, de forma que, na simulação de sistemas genéticos, a geração de números aleatórios adquire fundamental importância. Esses números são gerados em sequência que só se repete após um ciclo superior a ordem de trilhões de algoritmos. Dessa forma, os números gerados são pseudo-aleatórios. Para fins práticos, uma vez que a sequência completa jamais será usada, esses números comportam-se como números aleatórios (FERREIRA, 2001).

### 2.8.1 Uso da simulação de dados no melhoramento genético

Existe na literatura grande número de trabalhos utilizando simulação de dados, tanto na área vegetal como na área animal. A simulação pode ajudar a elucidar problemas em vários campos do melhoramento genético vegetal. Os primeiros trabalhos publicados nessa área foram apresentados por Fraser (1957a, 1957b). Esses estudos avaliaram o efeito da ligação nas taxas de ganho com a seleção massal e nos avanços dos ganhos genéticos.

Na área de delineamentos experimentais, a simulação tem sido amplamente utilizada. Silva et al. (2000) avaliaram a eficiência de quatro alternativas de experimentos em látice quanto à precisão na estimativa de componentes de variância. Os autores concluíram que o modelo de análise intrablocos deve ser utilizado em experimentos em látice para estimar componentes de variância sempre que a eficiência relativa do delineamento em látice for superior a 100% e, em caso contrário, deve-se optar pelo modelo de análise em blocos casualizados completos.

Tomé (2002) avaliou efeitos associados ao controle local nos delineamentos em látice para saber se estes devem ser analisados como fixos ou aleatórios. Por meio de simulação, verificou que se o tratamento for considerado fixo, o efeito de bloco é melhor representado como aleatório. Caso o efeito de tratamento for considerado aleatório, apenas com superfícies muito homogêneas e com herdabilidade alta, o controle local será melhor representado como efeito aleatório.

Vários trabalhos têm sido realizados no intuito de saber qual o melhor estimador de componentes de variância. Carneiro Junior et al. (2004), utilizando dados simulados de animais, compararam os métodos REML, ML e III de Henderson quanto ao efeito de desbalanceamento dos dados e da herdabilidade nas estimativas dos componentes de variância. Os autores verificaram que para características de alta herdabilidade não há diferença nos componentes de variância entre os métodos testados. Já para características com baixa herdabilidade, o método REML, com o modelo animal, pode ser considerado o mais apropriado para estimar componentes de variância.

Coelho (2005) comparou os métodos de quadrados mínimos e REML para obter estimativas de componentes de variância e concluiu que não houve diferença significativa entre os dois métodos testados para componentes de variância referente ao erro experimental, sendo que ambos estimaram componentes de variância próximos dos valores estabelecidos como referência. Já, para componente de variância referente ao efeito de tratamento, houve diferença marcante entre os dois métodos.

Assis et al. (2007) verificaram a influência de genes de efeito principal e do tamanho da população na estimação de componentes de variância genética aditiva, comparando-se as metodologias Bayesiana e REML, e concluíram que as duas metodologias, em geral, produzem resultados semelhantes na estimação dos componentes de variância.

A simulação também tem sido utilizada em trabalhos para obtenção de cruzamentos dialélicos e para responder questionamentos quanto à eficiência e contribuição dos marcadores moleculares no melhoramento de plantas. Entre os trabalhos realizados, Ferreira (1995) avaliou a eficiência do mapeamento de controladores de características quantitativas (QTLs) e da seleção assistida por marcadores.

Apesar da marcante contribuição da simulação de dados nos estudos de genética e melhoramento, seu emprego ainda é limitado. Essa limitação está associada ao fato de que profissionais da área de informática têm pouco conhecimento dos problemas da área de genética, os quais apresentam certa complexidade, por tratarem de fenômenos biológicos e envolverem princípios e complexas distribuições probabilísticas na sua análise. Por outro lado, são raros os geneticistas com conhecimento e aptidão para atuarem na área da informática (CRUZ, 2001).

### 3 MATERIAL E MÉTODOS

#### 3.1 SIMULAÇÃO DOS DADOS

O presente estudo foi realizado com dados simulados, utilizando-se o sistema computacional SAS<sup>®</sup> (SAS INSTITUTE, 2003). Foram simuladas populações, conforme método de simulação de Monte Carlo (HEUVELINK, 1998 citado por ORTIZ, 2004). A simulação dos valores fenotípicos baseou-se em dados reais obtidos a partir de experimentos de avaliação agrônômica de amendoim forrageiro (*Arachis pintoi* e *Arachis repens* Handro) conduzidos dentro do Programa de Melhoramento do Amendoim Forrageiro, coordenado pela Embrapa Acre (ASSIS; VALENTIM, 2009). A característica escolhida para análise foi produção de matéria seca (kg/ha) a cada corte.

Foram simuladas oito populações, considerando quatro tamanhos (número de indivíduos) e dois níveis de desbalanceamento (Tabela 1). A herdabilidade simulada foi de aproximadamente 0,30, valor este estimado para a característica em experimentos de campo. Considerou-se ausência de parentesco entre os indivíduos da mesma população, uma vez que não havia relação de parentesco entre os indivíduos nos experimentos utilizados como base para simulação. Para cada estrutura de população, o processo de simulação foi repetido 10 vezes, visando a validação do modelo estatístico usado na simulação e uma maior confiabilidade dos resultados obtidos.

Tabela 1 - Populações simuladas pelo método Monte Carlo, considerando diferentes tamanhos e níveis de desbalanceamento

População	Número de Indivíduos	Nível de Desbalanceamento (%)	Herdabilidade Simulada	Número Total de Observações (n)
1	4	0	0,30	192
2	4	20	0,30	154
3	10	0	0,30	480
4	10	20	0,30	384
5	30	0	0,31	1.440
6	30	20	0,31	1.152
7	100	0	0,29	4.800
8	100	20	0,29	3.840

O delineamento experimental utilizado foi em blocos casualizados completos, com seis repetições, considerando avaliações ao longo de dois anos consecutivos, com oito cortes.

Considerou-se o seguinte modelo estatístico para simulação dos dados:

$$y_{ijk} = \mu + g_i + b_j + ea_{ij} + c_k + gc_{ik} + bc_{jk} + e_{ijk}$$

$y_{ijk}$  é o valor observado do genótipo  $i$ , no bloco  $j$ , no corte  $k$ ;

$\mu$  é a média geral das observações;

$g_i$  é o efeito do genótipo  $i$ ;

$b_j$  é o efeito de bloco  $j$ ;

$ea_{ij}$  é o erro associado ao genótipo  $i$  no bloco  $j$ ;

$c_k$  é o efeito do corte  $k$ ;

$gc_{ik}$  é o efeito da interação do genótipo  $i$  no corte  $k$ ;

$bc_{jk}$  é o efeito da interação do bloco  $j$  no corte  $k$ ;

$e_{ijk}$  é o erro associado ao genótipo  $i$ , no bloco  $j$ , no corte  $k$ .

A média geral utilizada foi de 1.900 kg/ha de matéria seca produzida a cada corte, conforme valores encontrados em dados reais (ASSIS; VALENTIM, 2009). Os efeitos de genótipo, corte e bloco e as suas interações foram simulados de forma aleatória, considerando distribuição normal. A definição dos desvios padrão dos efeitos (Tabela 2) foi baseada nos valores de, herdabilidade e de repetibilidade encontrados usualmente nos dados oriundos de experimentos de campo. Dentro de uma mesma população, o efeito de genótipo simulado se manteve constante em todas as repetições realizadas.

Tabela 2 - Desvios padrão simulados para os diferentes tamanhos de populações

*Desvio padrão	POPULAÇÃO			
	4	10	30	100
$\sigma_g$	230	280	290	330
$\sigma_b$	30	30	30	30
$\sigma_c$	20	20	20	20
$\sigma_{bc}$	5	5	5	5
$\sigma_{gb}$	3	3	3	3
$\sigma_{gc}$	250	230	180	230
$\sigma$	730	650	600	650

\*Desvio padrão de: genótipo ( $\sigma_g$ ), bloco ( $\sigma_b$ ), corte ( $\sigma_c$ ), bloco x corte ( $\sigma_{bc}$ ), genótipo x bloco ( $\sigma_{gb}$ ), interação genótipo x corte ( $\sigma_{gc}$ ) e resíduo ( $\sigma$ )

Na simulação, foi considerado que a covariância entre duas observações decresce à medida que aumenta o intervalo de tempo entre elas. Após simulado cada efeito do modelo, os mesmos foram somados à média geral para obtenção dos valores fenotípicos ( $y_{ijk}$ ) simulados. Esses valores foram utilizados na análise estatística dos dados.

## 3.2 ANÁLISE DOS DADOS

### 3.2.1 Teste de esfericidade

Foi verificada a validade da condição de esfericidade da matriz de covariâncias residual entre cortes, através do teste de esfericidade de Mauchly (1940), descrito por Xavier (2000).

Após o teste de esfericidade, os dados simulados foram submetidos à análise de variância (ANOVA) pela metodologia de quadrados mínimos (STEEL; TORRIE, 1980) e à análise pela metodologia de modelos mistos (RESENDE, 2007).

### 3.2.2 Método de Quadrados Mínimos (LS)

O seguinte modelo estatístico foi considerado:

$$y_{ijk} = \mu + g_i + b_j + ea_{ij} + c_k + gc_{ik} + bc_{jk} + e_{ijk}$$

$y_{ijk}$  é o valor observado do  $i$ -ésimo genótipo no  $j$ -ésimo bloco no  $k$ -ésimo corte;

$\mu$  é a média geral das observações ;

$g_i$  é o efeito fixo do  $i$ -ésimo genótipo;

$b_j$  é o efeito fixo do  $j$ -ésimo bloco;

$ea_{ij}$  é o efeito do erro associado ao  $i$ -ésimo genótipo no  $j$ -ésimo bloco;

$c_k$  é o efeito fixo do  $k$ -ésimo corte;

$gc_{ik}$  é o efeito fixo da interação do  $i$ -ésimo genótipo com o  $k$ -ésimo corte;

$bc_{jk}$  é o efeito aleatório da interação do  $j$ -ésimo bloco no  $k$ -ésimo corte;

$e_{ijk}$  é o efeito do erro aleatório associado ao  $i$ -ésimo genótipo no  $j$ -ésimo bloco, no  $k$ -ésimo corte.

Foi realizada análise de variância, utilizando-se o comando PROC GLM do SAS® (SAS INSTITUTE, 2003). O delineamento foi em blocos casualizados completos em esquema de parcelas subdivididas

Adicionalmente, foram obtidas estimativas para os seguintes parâmetros: componentes quadráticos de genótipo, da interação genótipo x corte, da interação genótipo x bloco e componente de variância residual, por meio da igualdade dos quadrados médios a suas respectivas esperanças de quadrados.

O coeficiente de determinação genotípico ( $h_{gf}^2$ ) foi estimado de acordo com a expressão:

$$h_{gf}^2 = \frac{\phi_g^2}{\phi_g^2 + \phi_{gb}^2 + \phi_{gc}^2 + \hat{\sigma}_e^2}, \text{ onde:}$$

$\phi_g^2$  é o componente quadrático de genótipo;

$\phi_{gb}^2$  é o componente quadrático da interação genótipo x bloco;

$\phi_{gc}^2$  é o componente quadrático da interação genótipo x corte;

$\hat{\sigma}_e^2$  é o componente de variância residual.

Para identificar os indivíduos superiores, foi realizado o teste de Tukey a 5% de probabilidade. Apesar deste teste ser recomendado para populações de até 5 indivíduos (RESENDE; DUARTE, 2007), o mesmo foi empregado também para as populações com 10 e 30 genótipos, pelo fato deste procedimento ser bastante utilizado pelos melhoristas de plantas (BOTREL et al., 2000; SHIMOYA et al., 2002; FERREIRA et al., 1999). Para as populações de 100 indivíduos, no entanto, foram selecionados 15% dos melhores genótipos, utilizando a média fenotípica.

### 3.2.3 Metodologia de Modelos Mistos

O seguinte modelo estatístico foi utilizado:

$$y_{ijk} = \mu + g_i + b_j + c_k + gb_{ij} + gc_{ik} + bc_{jk} + e_{ijk}$$

$y_{ijk}$  é o valor observado do i-ésimo genótipo no j-ésimo bloco, no k-ésimo corte;

$\mu$  é a média geral das observações;  
 $g_i$  é o efeito aleatório de i-ésimo genótipo;  
 $c_k$  é o efeito fixo do k-ésimo corte ;  
 $b_j$  é o efeito fixo do j-ésimo bloco;  
 $gb_{ij}$  é o efeito aleatório da interação do i-ésimo genótipo com j-ésimo bloco.

$gc_{ik}$  é o efeito aleatório da interação do i-ésimo genótipo com k-ésimo corte;

$bc_{jk}$  é o efeito fixo da interação do j-ésimo bloco j no k-ésimo corte;

$e_{ijk}$  é o erro aleatório associado ao i-ésimo genótipo, no j-ésimo bloco, no k-ésimo corte.

Na forma matricial, tem-se:

$$y = Xb + Za + Wc + Tp + e, \text{ onde:}$$

$b$  é o vetor de dados dos efeitos fixos de bloco, corte e interação bloco x corte, somados a média geral;

$a$  é o vetor de dados dos efeitos aleatórios de genótipo;

$c$  é vetor de dados dos efeitos aleatórios de parcela;

$p$  é o vetor de dados do efeito aleatório de ambiente permanente;

$e$  é o vetor de erros aleatórios;

$X$ ,  $Z$ ,  $W$  e  $T$  são matrizes de incidência para os vetores  $b$ ,  $a$ ,  $c$  e  $p$ , respectivamente.

A estimação dos componentes de variância foi realizada pelo Método da Máxima Verossimilhança Restrita (REML) e, para predição dos valores genotípicos, empregou-se o método do Melhor Preditor Linear Não Viesado (BLUP). Os dados foram analisados utilizando-se o programa SELEGEN REML/BLUP (RESENDE, 2002b).

Foi estimada a herdabilidade individual no sentido amplo ( $h^2_g$ ) para as diferentes populações pela seguinte fórmula:

$$h_g^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_{gb}^2 + \hat{\sigma}_{gc}^2 + \hat{\sigma}_e^2}, \text{ em que:}$$

$\hat{\sigma}_g^2$  é o componente de variância genotípica;

$\hat{\sigma}_{gb}^2$  é o componente de variância da interação genótipo x bloco;

$\hat{\sigma}_{gc}^2$  é o componente de variância da interação genótipo x corte;

$\hat{\sigma}_e^2$  é o componente de variância residual.

### 3.2.4 Avaliação genética sob diferentes números de cortes

Com o intuito de identificar o número de cortes ideal para predição dos valores genotípicos dos indivíduos com alta confiabilidade, foram estimados os componentes de variância, as herdabilidades e as repetibilidades para as populações balanceadas, considerando a realização de 3, 4, 5, 6, 7 ou 8 cortes. A repetibilidade (r) foi determinada pela seguinte expressão:

$$r = \frac{\sigma_g^2 + \sigma_{gxb}^2}{\sigma_g^2 + \sigma_{gxb}^2 + \sigma_{gxc}^2 + \sigma^2}, \text{ em que:}$$

$\sigma_g^2$  é o componente de variância genotípica;

$\sigma_{gxb}^2$  é o componente de variância da interação genótipo x bloco;

$\sigma_{gxc}^2$  é o componente de variância da interação genótipo x corte;

$\sigma^2$  é o componente de variância residual;

O coeficiente de determinação ( $R^2$ ) foi calculado pela seguinte expressão:

$$R^2 = \frac{nr}{1 + r(n - 1)}, \text{ em que:}$$

n é o número de cortes;

r é a repetibilidade da característica.

Utilizou-se o programa SELEGEN para obtenção de tais estimativas e da predição dos valores genotípicos.

### 3.3 CRITÉRIOS DE AVALIAÇÃO

Foram obtidas as médias e os desvios padrão das estimativas de componentes de variância, herdabilidade individual no sentido amplo, componentes quadráticos e coeficientes de determinação genotípico,

estimados pelos métodos REML e ANOVA, conforme os modelos apresentados, para cada uma das populações simuladas, a partir das 10 repetições realizadas.

### 3.3.1 Acurácia das estimativas de parâmetros genéticos

Para avaliação dos valores dos componentes de variância e herdabilidades estimados pelo REML e dos componentes quadráticos e coeficientes de determinação genotípico estimados pela ANOVA, calculou-se a porcentagem de erro (PE), conforme Carneiro Junior (2005):

$$PE = \frac{(\hat{E} - R)100}{R}, \text{ em que:}$$

$\hat{E}$  é o valor estimado e R é o valor real do respectivo parâmetro

Por meio da porcentagem de erro da estimação dos componentes de variância, componentes quadráticos, herdabilidades e coeficientes de determinação genotípico, é possível verificar qual metodologia entre as estudadas estima parâmetros genéticos com maior ou menor erro em relação ao valor real (simulado). Quanto mais próximo de zero mais acurada é a estimativa.

### 3.3.2 Correlação entre valores genotípicos

Após a predição dos valores genotípicos pelo método REML/BLUP, foi calculada a correlação de Spearman entre valores genotípicos preditos e reais. Esta correlação mede a intensidade da relação entre duas variáveis ordinais, utilizando, em vez do valor observado, apenas a ordem das observações.

As correlações de Spearman entre valores genotípicos preditos e reais foram obtidas pelo SAS<sup>®</sup> (SAS INSTITUTE, 2003), conforme Cruz (2001), pela seguinte expressão:

$$CS = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \text{ em que:}$$

$d_i$  é a diferença de posto para cada par de observações;

$n$  é o número de postos.

#### 3.3.4 Quadrado Médio do Erro (QME)

Os valores genotípicos preditos pelo método REML/BLUP foram comparados com os valores genotípicos reais, por meio quadrado médio do erro (QME), como em Harvey (1979), citado por Schenkey (1998):

$$QME = \sum_{i=1}^n \frac{1}{n} (\hat{a}_i - a_i)^2, \text{ em que:}$$

$\hat{a}_i$  e  $a_i$  são os valores genotípicos preditos e reais do  $i$ -ésimo indivíduo, respectivamente;

$n$  é o número de indivíduos.

Quanto mais próximo de zero, maior é a acurácia da predição dos valores.

#### 3.3.5 Acurácia da predição dos valores genotípicos em função do número de cortes

Calcularam-se a correlação de Spearman e o Quadrado Médio do Erro (QME) para as populações balanceadas de 4, 10, 30 e 100 indivíduos, considerando a realização de 3, 4, 5, 6, 7 e 8 cortes, a partir das 10 repetições.

#### 3.3.6 Ranqueamento de genótipos e ganho de seleção

O ganho de seleção real obtido pelas duas metodologias foi calculado, uma vez que os valores genotípicos reais de cada indivíduo são conhecidos. Para a metodologia de quadrados mínimos, por meio do teste de médias Tukey, foram utilizados os genótipos que obtiveram letra A, para as populações de 4, 10 e 30 genótipos. Para a população de 100 genótipos, utilizou-se uma intensidade de seleção de 15%. Para a metodologia REML/BLUP, foram utilizadas intensidades de seleção de 25%, 20%, 15% e 15% para populações de 4, 10, 30 e 100 genótipos, respectivamente. Foram obtidas estatísticas descritivas (média e desvio padrão) dos valores genotípicos reais dos genótipos selecionados pelo teste de Tukey e ranqueados pelo BLUP, a partir das 10 repetições.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 TESTE DE ESFERICIDADE

Os dados analisados apresentaram resultados significativos para o teste de esfericidade de Mauchly para todas as populações estudadas, balanceadas (Tabela 3) e desbalanceadas (Tabela 4). Este resultado confirma que a matriz de covariância dos dados em estudo não é homogênea. A rigor, a análise dos dados não poderia ser realizada considerando o delineamento em parcelas subdividas no tempo (FREITAS, 2008), que leva em consideração que a covariância entre as medidas repetidas são homogêneas. Os níveis do fator tempo alocados nas subparcelas nesse tipo de delineamento não permitem a aleatorização e, com isso, a matriz de covariâncias não apresenta estrutura igual à exigida para que a análise usual univariada seja válida.

Tabela 3 - Teste de esfericidade de Mauchly (W) para diferentes tamanhos de populações balanceadas, com base no teste de Qui-quadrado ( $\chi^2$ )

Tamanho da População	W	$\chi^2$	Pr.> $\chi^2$
4	$6,39 \times 10^{-3}$	613,48	<0,0001
10	$3,97 \times 10^{-3}$	1.678,57	<0,0001
30	$6,35 \times 10^{-5}$	8.360,64	<0,0001
100	$1,54 \times 10^{-9}$	58.726,63	<0,0001

Tabela 4 - Teste de esfericidade de Mauchly (W) para diferentes tamanhos de populações desbalanceadas, com base no teste de Qui-quadrado ( $\chi^2$ )

Tamanho da População	W	$\chi^2$	Pr.> $\chi^2$
4	0,0284	389,31	<0,0001
10	0,0441	3.622,08	<0,0001
30	0,0150	886,15	<0,0001
100	$1,144 \times 10^{-5}$	32.915,68	<0,0001

Xavier e Dias (2001), em estudo com simulação de dados, analisaram a acurácia da análise em modelos univariado e multivariado em casos onde a matriz de covariância atendia, ou não, a condição de esfericidade e concluíram que a utilização de matrizes de covariâncias que não atendem a condição de esfericidade leva a resultados inválidos para o teste F dos fatores das subparcelas. Os mesmos autores ressaltam que uma alternativa para o uso

desta metodologia seria o ajuste do número de graus de liberdade do teste F para o fator da subparcela.

Em estudos realizados por Malheiros (2004) com dados simulados, o autor concluiu que os testes F são precisos quando se faz correção do número de graus de liberdade do fator tempo, dependendo do método de correção utilizado. Em trabalhos com avaliação e seleção de plantas forrageiras tropicais, a correção dos dados para o fator da subparcela geralmente não é realizado, sendo que esse efeito pode causar seleção de genótipos não superiores.

Freitas et al. (2011) salientam que a análise de dados de plantas forrageiras tem sido realizada de maneira incorreta com o uso de delineamento em esquema de parcelas subdivididas. Nesse contexto, os autores realizaram um estudo com a cultura da alfafa testando o modelo linear geral (quadrados mínimos) e modelo misto, avaliando diferenças entre os procedimentos GLM e MIXED do SAS<sup>®</sup> (SAS INSTITUTE, 2003) em análises com medidas repetidas. Os autores concluíram que as análises com medidas repetidas utilizando quadrados mínimos somente produz resultados corretos quando a condição de esfericidade da matriz de covariância residual entre cortes é atendida.

Andrade (2013), em estudo analisando as implicações da modelagem estatística na seleção de progênies de cafeeiro, realizou análise de correlação de Spearman entre as colheitas e verificou correlação alta entre estas. O autor também encontrou resultado do teste de esfericidade de Mauchly significativo, confirmando que em culturas perenes a condição de esfericidade da matriz de covariância geralmente não é satisfeita. Este autor recomenda que para análise de dados com medidas repetidas deve ser feito um estudo para verificação de qual matriz de covariância mais se adequa aos dados.

Ao não atender a condição de esfericidade, a taxa de erro tipo I aumenta, ou seja, a probabilidade de rejeitar a hipótese nula quando esta é verdadeira é maior. Esse erro pode levar a inferências equivocadas, levando o melhorista a selecionar uma população geneticamente não variável, impedindo o ganho com a seleção.

Observa-se na literatura (DAHER et al., 2004; FERREIRA et al., 2012; LÁZARO et al., 2009; MARTUSCELLO et al., 2007) que este fato não é levado em consideração na análise de dados de forrageiras tropicais, quando estes são analisados pela metodologia de quadrados mínimos. A utilização da

metodologia de modelos mistos seria uma alternativa viável, tendo em vista que esta metodologia tem a vantagem de permitir a escolha da matriz que mais se adéqua aos dados da cultura avaliada.

## 4.2 ESTIMATIVAS DE PARÂMETROS GENÉTICOS

As médias das estimativas dos componentes de variância e componentes quadráticos obtidas pelos métodos REML e ANOVA, respectivamente, foram, em geral, similares, tanto para os dados balanceados quanto para os dados desbalanceados (Tabela 5 e Tabela 6).

### 4.2.1 Variação do genótipo

A média dos componentes de variância genotípica ( $\sigma^2_g$ ) estimados pelo REML, para população desbalanceada com quatro genótipos, apresentou menor desvio padrão, em comparação aos valores estimados pela ANOVA para a mesma população (Tabela 6) Este resultado ocorreu devido ao método REML estimar valores mais próximos à média das 10 repetições nas populações desbalanceadas. Para a população balanceada com quatro genótipos as metodologias REML e ANOVA apresentaram desvios padrão similares (Tabela 5 e Tabela 6).

A estimativa do componente de variância genotípica ( $\sigma^2_g$ ) pelo REML e do componente quadrático de genótipo ( $\phi_g^2$ ) estimado pela ANOVA nas populações balanceada e desbalanceada com quatro genótipos foram superestimadas (Figura 2). As duas metodologias apresentaram porcentagem de erro na estimação da variância genotípica e componente quadrático de genótipo de aproximadamente 72% para a população balanceada com quatro indivíduos. Para a população desbalanceada o método ANOVA apresentou maior porcentagem de erro, aproximadamente 90%, em comparação com o método REML, que apresentou porcentagem de erro em torno de 80% (Figura 2).

Em comparação com as populações de 10, 30 e 100 genótipos, a população com quatro genótipos apresentou maior porcentagem de erro na estimação da variância genotípica e do componente quadrático de genótipo pelas metodologias REML e ANOVA, evidenciando que quanto maior o número de indivíduos, mais acurada é a estimativa de parâmetros genéticos (Figura 2).

Tabela 5 - Média e desvio padrão (DP) das estimativas de parâmetros genéticos obtidas pelos métodos REML e ANOVA para diferentes tamanhos de populações balanceadas

Método	NÚMERO DE GENÓTIPOS							
	4		10		30		100	
ANOVA <sup>1</sup>	Média	DP	Média	DP	Média	DP	Média	DP
$\Phi^2_g$	90.693,68	37.249,75	99.382,78	19.666,97	84.168,93	11.882,9	102.680,50	9.480,04
$\Phi^2_{gxb}$	47.251,12	28.223,16	58.642,20	11.178,55	39.370,35	7.592,99	42.098,58	3.477,73
$\Phi^2_{gxc}$	58.090,67	18.572,46	49.332,24	12.844,43	38.420,12	5.002,59	53.084,46	2.376,42
$\sigma^2$	174.661,20	23.062,83	173.538,20	9.155,32	123.584,00	6.635,44	144.636,40	5.294,59
$h^2_{gf}$	0,24	0,0810	0,26	0,0517	0,29	0,0391	0,30	0,0216
REML <sup>2</sup>								
$\sigma^2_g$	90.695,74	37.246,71	99.831,00	19.303,69	84.170,82	11.883,04	102.682,90	9.480,06
$\sigma^2_{gxb}$	47.250,16	28.214,43	58.110,83	11.112,51	39.366,31	7.594,54	42.092,80	3.478,46
$\sigma^2_{gxc}$	58.080,96	18.577,06	49.107,84	11.654,44	38.413,28	5.003,54	53.077,36	2.376,92
$\sigma^2$	174.665,60	23.066,89	172.382,40	9.703,35	123.589,80	6.636,03	144.643,0	5.295,08
$h^2_g$	0,24	0,0809	0,26	0,0507	0,29	0,0391	0,30	0,0215

<sup>1</sup> $\Phi^2_g$ : componente quadrático de genótipo;  $\Phi^2_{gxb}$ : componente quadrático da interação genótipo x bloco;  $\Phi^2_{gxc}$ : componente quadrático da interação genótipo x corte;  $\sigma^2$ : componente de variância residual;  $h^2_{gf}$ : coeficiente de determinação genotípico.

<sup>2</sup> $\sigma^2_g$ : variância genotípica;  $\sigma^2_{gxb}$ : variância da interação genótipo bloco;  $\sigma^2_{gxc}$ : variância da interação genótipo x corte;  $\sigma^2$ : componente de variância residual;  $h^2_g$ : herdabilidade.

Tabela 6 - Média e desvio padrão (DP) das estimativas de parâmetros genéticos obtidas pelos métodos REML e ANOVA para diferentes tamanhos de populações desbalanceadas

Método	NÚMERO DE GENÓTIPOS							
	4		10		30		100	
ANOVA <sup>1</sup>	Média	DP	Média	DP	Média	DP	Média	DP
$\Phi^2_g$	99.295,48	47.574,65	102.069,90	25.057,42	84.035,41	11.193,17	102.407,50	10.418,52
$\Phi^2_{gxb}$	43.529,99	25.159,40	57.036,50	12.959,78	38.920,55	10.580,94	41.126,53	4.422,36
$\Phi^2_{gxc}$	52.043,83	27.513,83	50.184,89	15.004,42	39.106,25	6.107,54	52.978,07	2.662,95
$\sigma^2$	185.982,60	33.048,52	180.669,80	12.239,84	124.907,60	8.177,42	142.659,20	5.051,27
$h^2_{gf}$	0,25	0,1003	0,26	0,0614	0,29	0,0414	0,30	0,0247
REML <sup>2</sup>								
$\sigma^2_g$	94.316,93	44.520,40	100.429,40	23.754,14	83.406,20	10.798,32	102.219,80	9.784,02
$\sigma^2_{gxb}$	44.044,63	27.064,95	56.983,41	13.192,93	39.377,81	10.619,03	41.104,22	3.847,99
$\sigma^2_{gxc}$	52.184,08	24.504,29	50.339,98	13.458,70	38.505,93	6.123,53	53.290,44	2.885,44
$\sigma^2$	184.797,10	32.996,18	179.104,60	14.229,30	125.083,50	8.193,37	142.733,70	4.986,84
$h^2_g$	0,25	0,0976	0,26	0,0603	0,29	0,0406	0,30	0,0232

<sup>1</sup> $\Phi^2_g$ : componente quadrático de genótipo;  $\Phi^2_{gxb}$ : componente quadrático da interação genótipo x bloco;  $\Phi^2_{gxc}$ : componente quadrático da interação genótipo x corte;  $\sigma^2$ : componente de variância residual;  $h^2_{gf}$ : coeficiente de determinação genotípico.

<sup>2</sup> $\sigma^2_g$ : variância genotípica;  $\sigma^2_{gxb}$ : variância da interação genótipo bloco;  $\sigma^2_{gxc}$ : variância da interação genótipo x corte;  $\sigma^2$ : componente de variância residual;  $h^2_g$ : herdabilidade.

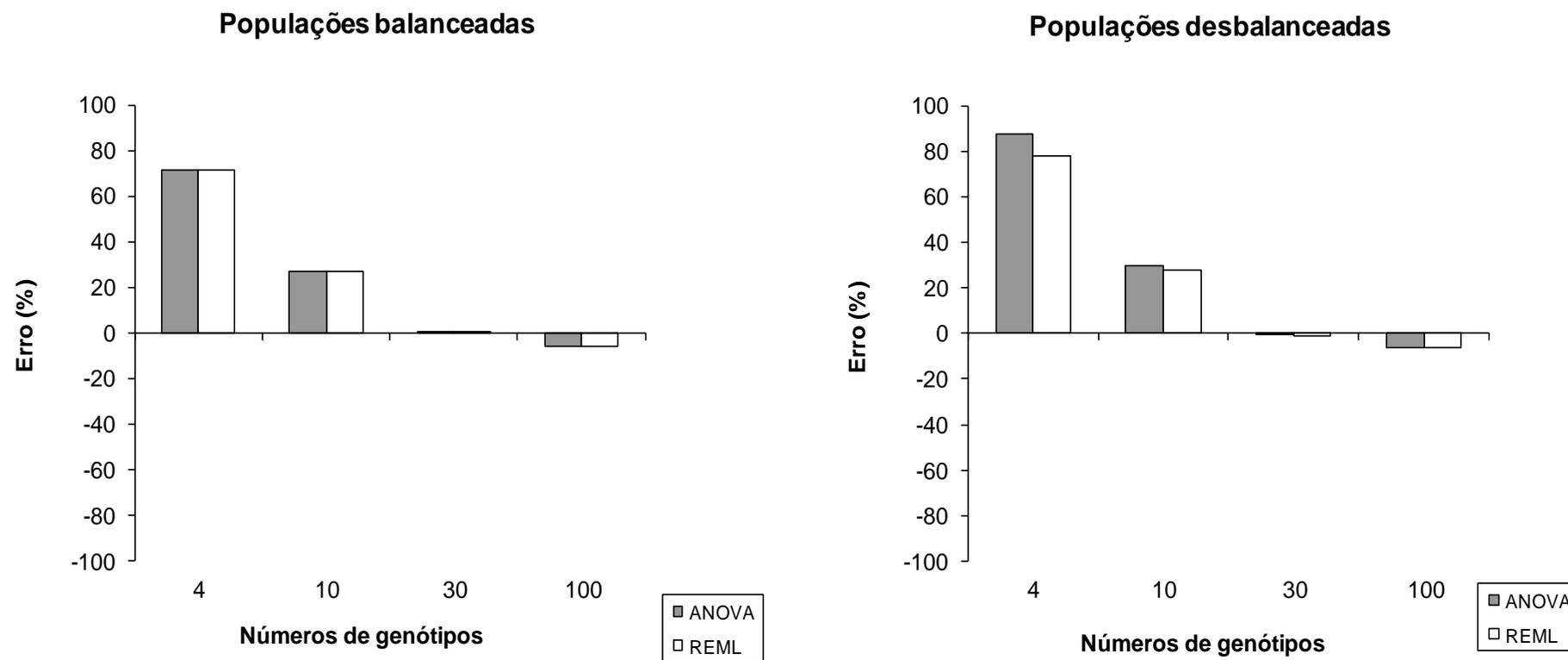


Figura 2 - Porcentagem de erro das estimativas dos componentes de variância e componentes quadráticos de genótipo, obtidos pela metodologia REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas

Para a população com 10 genótipos os dois métodos utilizados superestimaram o resultado. Na população balanceada, as estimativas dos componentes de variância e componentes quadráticos de genótipos foram similares (Figura 2), porém a metodologia REML apresentou, em geral, menor desvio padrão da média da estimativa do componente de variância genotípico, dando uma maior confiabilidade ao resultado (Tabela 5). Segundo Viana et al. (2012) os métodos REML/BLUP e ANOVA apresentam resultados similares de componentes de variância quando os dados são balanceados.

Na população desbalanceada com 10 genótipos, apesar de as duas metodologias terem superestimado o valor da variância genotípica e do componente quadrático de genótipo, as estimativas obtidas pelo método REML também se aproximaram mais do valor real, com menor desvio padrão (Tabela 6). Os resultados de Farias Neto e Resende (2001) corroboram com o presente estudo. Os referidos autores, testando os métodos ANOVA e REML na estimação de componentes de variância em dados experimentais de pupunheira, mostraram que esses métodos levam a estimativas de componentes de variância divergentes em dados desbalanceados. O autor recomenda o método REML para estimação de componentes de variância de plantas perenes para esta situação.

Por outro lado, Silva et al. (2000) encontraram resultados diferentes avaliando quatro alternativas de experimento em látice com dados simulados, com baixa e alta eficiência do látice quanto à estimação de componentes de variância via quadrados mínimos (ANOVA). Os autores mostraram que a metodologia não apresenta vício, estimando com precisão a variância de tratamento em três dos quatro modelos testados. Porém, os dados analisados por Silva et al. (2000) não apresentavam medidas repetidas nem desbalanceamento de dados, o que pode ter ocasionado tais diferenças.

Quanto à porcentagem de erro (Figura 2), observa-se que os métodos ANOVA e REML para a população balanceada com 10 genótipos apresentaram resultados similares, com aproximadamente 27% de erro. Para a população desbalanceada, a metodologia REML apresentou porcentagem de erro levemente inferior (2,09%) ao método ANOVA.

Nas populações balanceadas e desbalanceadas com 30 e 100 genótipos, as estimativas obtidas pelas metodologias REML e ANOVA para componentes de variância e componente quadrático de genótipo foram

próximas ao valor real (Figura 2). Nessas populações, quando considerado desbalanceamento, os desvios padrão da média das estimativas de componentes de variância e componente quadrático de genótipo, foram menores em comparação aos desvios padrão apresentados pelo método ANOVA (Tabela 5 e Tabela 6).

Foi observado que o aumento do número de genótipos para 30 e 100, as estimativas dos componentes de variância e do componente quadrático de genótipo ficaram mais próximas do valor real, apresentando porcentagem de erro próximo de zero (Figura 2). Este resultado mostra que mesmo com desbalanceamento as duas metodologias geram estimativas acuradas quando há elevado número de observações que variaram de 1.152 a 4.800 (Tabela 1). Duarte (2000), em estudo com simulação de dados, testando diferentes métodos de estimativa de componentes de variância (REML, MIVQUE, ANOVA e ML), em delineamento em blocos aumentados, observou que, para experimentos com elevado número de observações ( $n > 200$ ), os métodos ANOVA e REML apresentam estimativas de boa qualidade. No presente trabalho, observaram-se menores porcentagens de erro (próximos de zero) quando as análises foram realizadas considerando número de observações de 1.152 a 4.800.

#### 4.2.2 Variação da interação genótipo x bloco

A variação do componente quadrático e componente de variância da interação genótipo x bloco, estimadas pelo método ANOVA e REML, foi superestimada em todas as populações estudadas (Tabela 5 e Tabela 6). O valor estimado pelas metodologias para esse efeito foi muito acima do valor real, com porcentagens de erro acima de 1.000%. Este resultado se deve, provavelmente, à complexidade do modelo estatístico simulado, que além dos efeitos de corte, bloco e genótipo, possui as suas devidas interações. Possivelmente, parte da variação residual, que foi subestimada, foi computada como variação do efeito da interação genótipo x bloco.

Em estudos com simulação esta situação não tem sido observada com frequência, porém os modelos simulados são menos parametrizados, como o de Coelho e Barbin (2006) que trabalharam com delineamento inteiramente casualizado e em blocos casualizados sem interações; Martinez (2010), que simulou valores fenotípicos de clones de espécie madeireira de acordo com

modelo em blocos casualizados sem considerar as interações entre os efeitos, não obtendo problemas quanto à estimativa das variâncias relacionadas ao modelo.

#### 4.2.3 Variação da interação genótipo x corte

As estimativas do componente de variância e do componente quadrático da interação genótipo x corte para as populações balanceada e desbalanceada de quatro e 10 genótipos foram levemente subestimadas (Figura 3). Para a população com quatro genótipos, as estimativas dos parâmetros referentes ao efeito da interação genótipo x corte obtidas pelos métodos ANOVA e REML foram mais próximas do valor real na população balanceada. Observou-se para esta população que a porcentagem de erro foi de apenas 7%, tanto para a ANOVA quanto para o REML, em comparação com a população desbalanceada que apresentou porcentagem de erro de aproximadamente 17% (Figura 3).

A população com quatro genótipos apresentou desvio padrão da média das estimativas de componente quadrático e componente de variância da interação genótipo x corte similar, considerando a população balanceada. Porém, para população desbalanceada as estimativas desse componente pelo método REML apresentou desvio padrão menor em comparação com o desvio das estimativas obtidas pela ANOVA (Tabela 5 e Tabela 6).

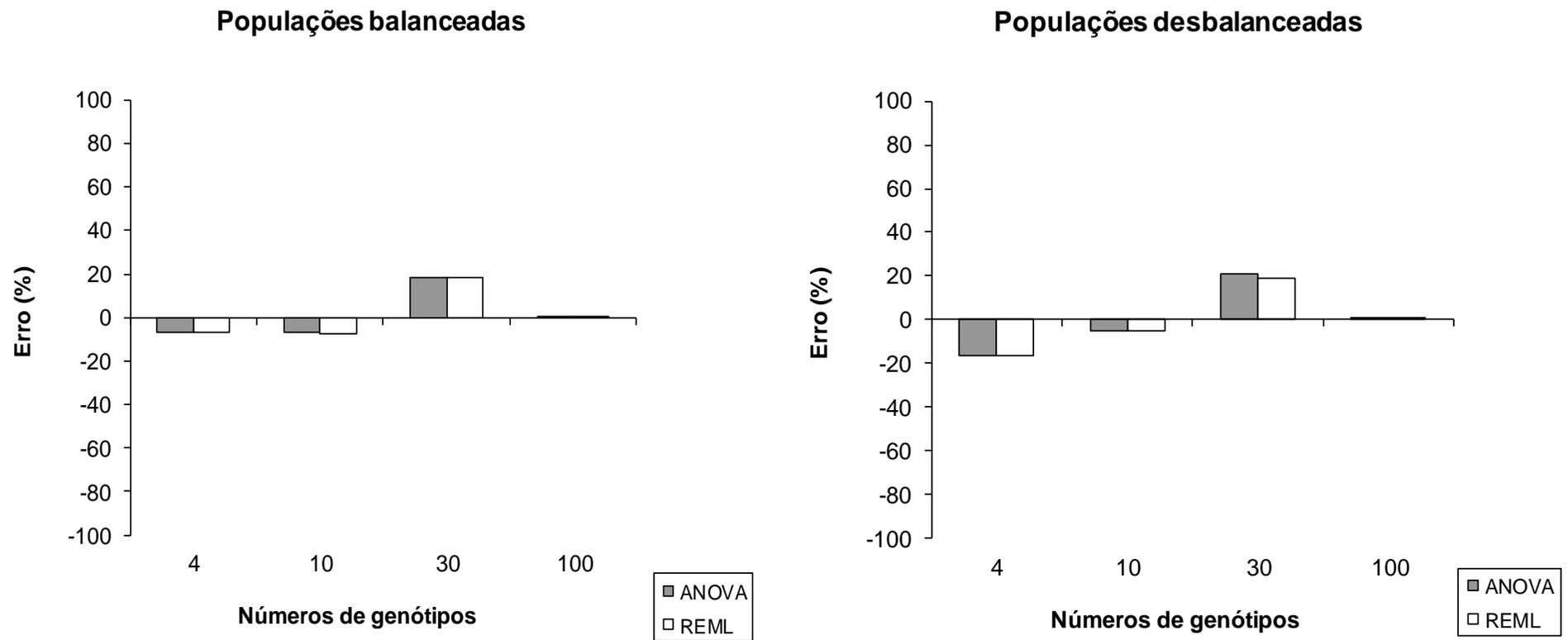


Figura 3 - Porcentagem de erro das estimativas dos componentes de variância e componentes quadráticos da interação genótipo x corte, obtidos pela metodologia REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas

Na população com 10 genótipos, as estimativas obtidas por ambas as metodologias, para as populações balanceada e desbalanceada foram próximas ao valor real, sendo levemente subestimadas (Figura 3), com porcentagem de erro para população balanceada e desbalanceada semelhantes, variando de 4% a 8% (Figura 3).

As populações balanceadas com quatro e 10 genótipos apresentaram menores desvios padrão para as estimativas do componente quadrático e componente de variância da interação genótipo x corte, em comparação com a população desbalanceada para o mesmo tamanho de população (Tabela 5 e Tabela 6). Este resultado mostra que ambas as metodologias estimaram estes efeitos com maior precisão quando a população analisada é balanceada, em que a dispersão dos valores estimados em relação à média é menor, considerando as 10 repetições.

Para as populações com 30 e 100 indivíduos, as metodologias REML e ANOVA superestimaram a variância e o componente quadrático da interação genótipo x corte nas populações balanceadas e desbalanceadas. Para a população balanceada com 30 genótipos, houve pequena variação do valor estimado por ambas as metodologias (Tabela 5 e Tabela 6). Na população desbalanceada, de mesmo tamanho, o método REML estimou valor mais próximo do valor real (Figura 3). Segundo Zeni-Neto et al. (2008), um dos principais problemas na seleção e na recomendação de genótipos de qualquer programa de melhoramento, é a interação genótipo x ambiente, e a utilização de modelos mistos na análise permite lidar com desbalanceamento de dados.

A porcentagem de erro das estimativas de componente de variância e componente quadrático da interação genótipo x corte obtida pelos métodos REML e ANOVA para população com 30 genótipos foi similar na população balanceada, sendo de aproximadamente 19%. Essa porcentagem de erro aumentou quando se consideram as estimativas desse componente feito pelo método ANOVA na população desbalanceada, sendo de aproximadamente 21% (Figura 3).

Os desvios padrão da média das estimativas feitas por ambas as metodologias, na população com 30 genótipos, foram similares, tanto na população balanceada quanto na população desbalanceada (Tabela 5 e Tabela 6). No entanto, a população desbalanceada apresentou maior desvio padrão da

média em comparação com a população balanceada para as estimativas feitas pelos métodos ANOVA e REML.

Na população com 100 genótipos, as estimativas referentes ao efeito da interação genótipo x corte das populações balanceada e desbalanceada obtidas pelos métodos REML e ANOVA foram similares, sendo bem próximos do valor real, de forma que as porcentagens de erro foram próximas de zero (Figura 3). Os desvios padrão observados foram relativamente de baixa magnitude, sendo maior na população desbalanceada (Tabela 5 e Tabela 6). Esses resultados evidenciam que quanto maior a população, maior a acurácia na estimativa dos parâmetros genéticos.

A porcentagem de erro da estimativa do componente quadrático e componente de variância da interação genótipo x corte foi baixa, sendo quase zero para população balanceada e desbalanceada (Figura 3).

#### 4.2.4 Variância residual

As duas metodologias subestimaram as estimativas de componentes variância residual ( $\sigma^2$ ), para as populações balanceada e desbalanceada, com porcentagens de erro variando de 58% à 67% (Figura 3). As metodologias REML e ANOVA estimaram componentes de variância residual similares nas populações balanceadas e desbalanceadas (Tabela 5 e Tabela 6).

Duarte et al. (2001), em trabalho com simulação de experimentos em delineamento em blocos casualizados aumentados, testando vários métodos de estimação de componentes de variância (ANOVA, REML, MIVQUE e ML) constataram que o método ANOVA mostrou-se não tendencioso na estimação da variância residual, porém apresentou grande flutuação nos valores médios desta estimativa. Segundo esses autores, o método REML apresentou componentes de variância residual subestimados quando o número de observações foi menor que 120 e o número de blocos menor que seis. Para experimentos maiores, com número de observações maior que 240 e número de blocos maior ou igual a 24, este método apresentou estimativa de variância residual de boa qualidade, com viés tendendo para zero.

No presente estudo, as estimativas dos componentes de variância residual obtidas pelos métodos ANOVA e REML foram subestimadas, mesmo nas populações maiores, com 100 genótipos e com número de observações igual 4.800.

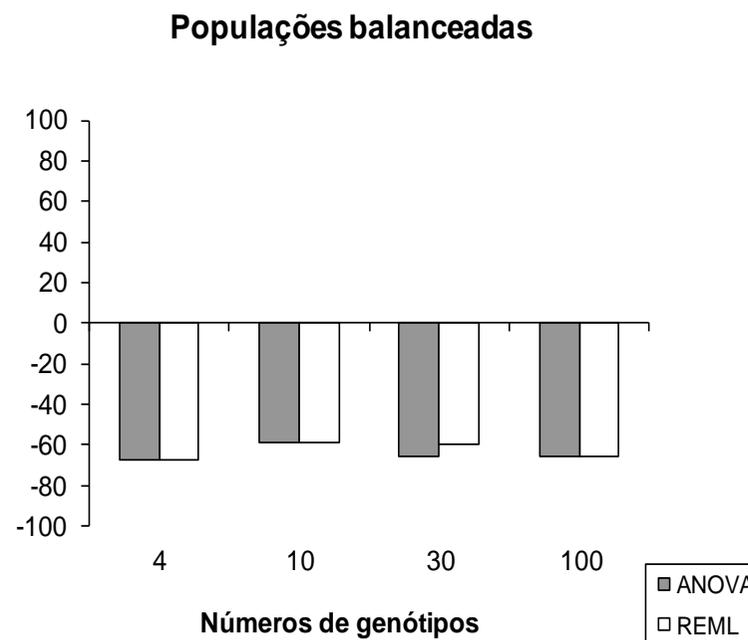
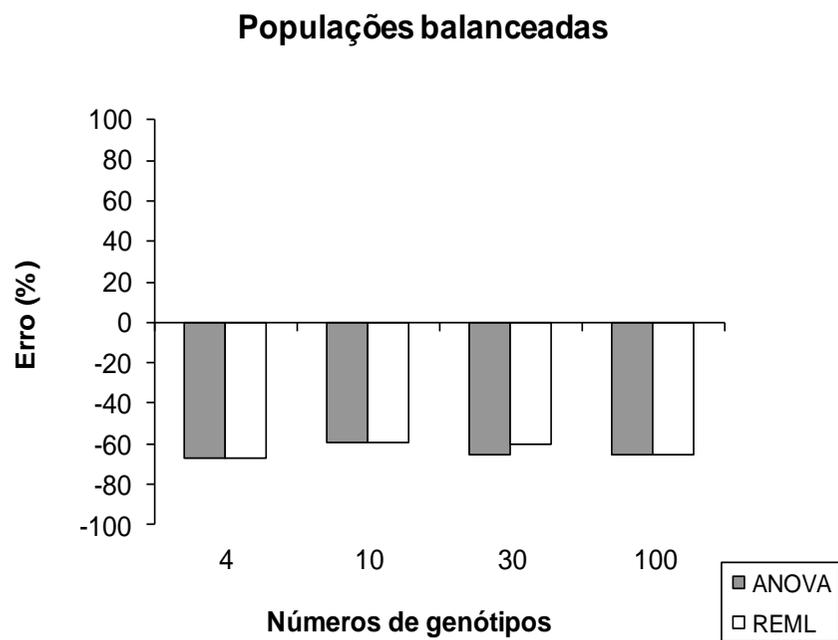


Figura 4 - Porcentagem de erro das estimativas dos componentes de variância residual, obtidos pelas metodologias REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas

Os desvios padrão da média das estimativas da variância residual obtidas pelos métodos ANOVA e REML foram similares em todas as populações. Porém, na população desbalanceada, os desvios padrão foram maiores em comparação com os desvios padrão apresentados pelas populações balanceadas, exceto para população com 100 genótipos, onde estes foram similares, tanto para população balanceada quando desbalanceada. Este resultado evidencia que nesta situação as populações com 100 genótipos balanceada e desbalanceada apresentam estimativas menos dispersas em relação à média.

#### 4.2.5 Herdabilidade e coeficiente de determinação genotípico

O conhecimento da herdabilidade e do coeficiente de determinação genotípico da característica estudada depende das estimativas dos componentes de variância e dos componentes quadráticos. Neste estudo, para as populações com 4, 10 e 30 genótipos as metodologias REML e ANOVA geraram estimativas subestimadas de herdabilidade e coeficiente de determinação genotípico, respectivamente, nas populações balanceadas e desbalanceadas (Figura 5).

Nas populações com 100 genótipos, as estimativas obtidas pelos métodos REML e ANOVA foram mais próximas do valor real, sendo levemente superestimadas (Figura 5). Observando-se as porcentagens de erro em relação à herdabilidade estimada pela metodologia REML e ao coeficiente de determinação genotípico estimado pelo método ANOVA, notou-se que quanto maior o tamanho da população, menor é o erro na estimativa desses parâmetros genéticos (Figura 5).

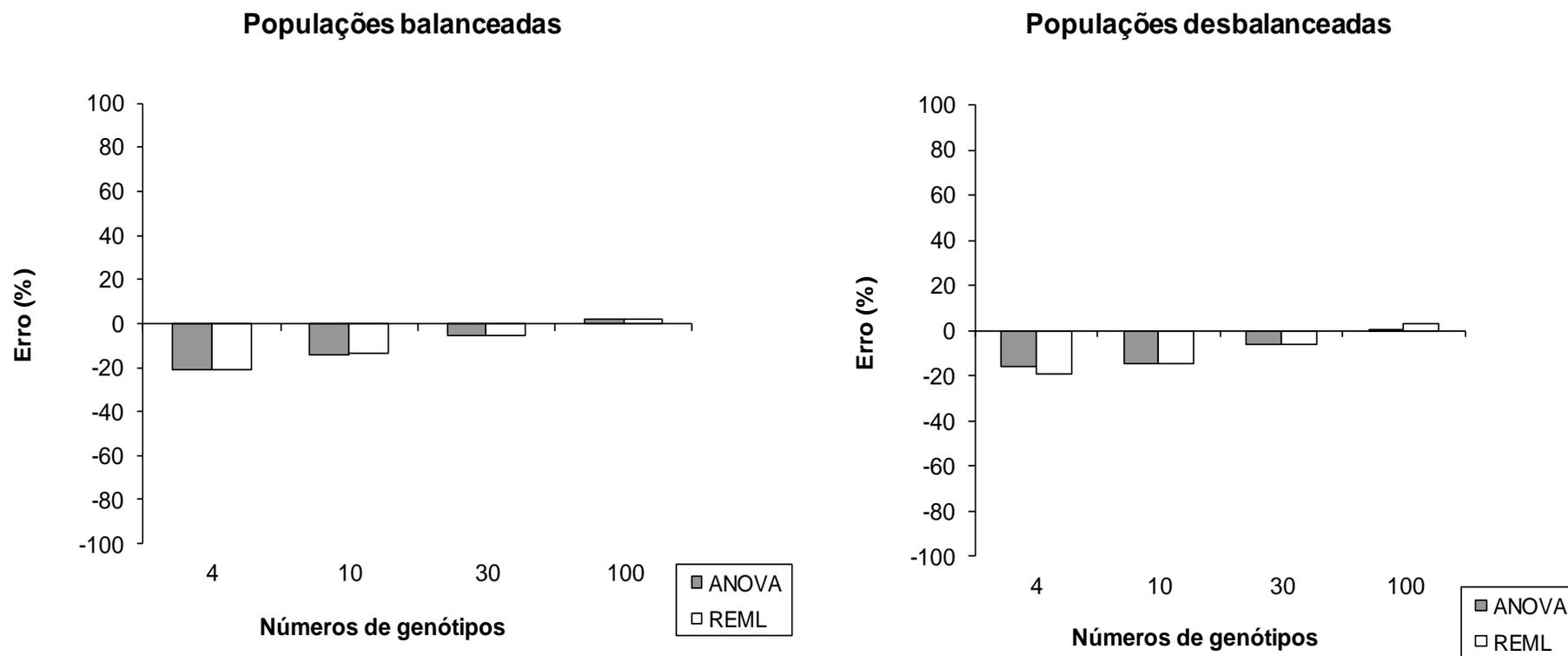


Figura 5 - Porcentagem de erro das estimativas de herdabilidade e do coeficiente de determinação genotípico, obtidos pelas metodologias REML e ANOVA em populações com diferentes tamanhos, balanceadas e desbalanceadas

As populações balanceada e desbalanceada com quatro genótipos apresentaram porcentagem de erro de aproximadamente 21% para as estimativas da herdabilidade e coeficiente de determinação genotípico. Nas populações com 10, 30 e 100 genótipos, a porcentagem de erro para as estimativas dos mesmos parâmetros citados anteriormente, tenderam a diminuir em populações balanceadas e desbalanceadas, sendo que para a população com 100 genótipos, as porcentagens de erro foram praticamente zero (Figura 5).

Estudos de simulação baseados em dados experimentais de milho realizados por Estaghvirou et al. (2013), que avaliaram a precisão de sete alternativas para obtenção da estimativa de herdabilidade considerando modelo misto de análise dos dados, concluíram que um maior conjunto de dados leva a uma maior precisão da estimativa da herdabilidade, corroborando com os resultados do presente estudo.

Apesar das metodologias ANOVA e REML subestimarem as estimativas do componente de variância residual para todas as amostras estudadas, este problema não interferiu de forma expressiva nas estimativas de herdabilidade para as populações balanceadas e desbalanceadas das diferentes populações estudadas. Porém, de maneira geral, as estimativas da herdabilidade e do coeficiente de determinação foram influenciadas pelo número de genótipos. Ao aumentar o tamanho da população, verificou-se que as estimativas desse parâmetro se aproximaram mais do valor real.

Piepho e Mohring (2007) relatam que os cálculos usuais para herdabilidade apresentados pelos modelos tradicionais não levam em consideração a correlação genotípica, são baseados na independência dos erros e blocos aleatórios e não se aplicam a dados desbalanceados. Os autores afirmam que o cálculo da herdabilidade em condições de dados desbalanceados e com heterogeneidade de variâncias prejudicam a inferência de resposta à seleção. No presente estudo, o desbalanceamento de dados e o uso do modelo tradicional de quadrados mínimos (ANOVA) para o cálculo da herdabilidade não causaram problemas na estimação desse parâmetro em populações com grande número de observações (> 1.000). Porém em populações menores, com número de observações menor que 500 (4 e 10 genótipos), observou-se menor acurácia na estimativa da herdabilidade para ambas as metodologias.

Coelho (2005), em estudo com simulação de dados, testou os métodos ANOVA e REML na estimação de herdabilidade e componentes de variância, em dois modelos de análise e diferentes herdabilidades. O autor concluiu que os métodos promovem estimativas diferentes quando a herdabilidade é inferior a 0,30; quando esta é de 0,30 a 0,70, os métodos REML e ANOVA estimam herdabilidades similares. Este resultado corrobora com o encontrado no presente trabalho.

#### 4.3 CORRELAÇÃO DE SPEARMAN E QUADRADO MÉDIO DO ERRO

A Correlação de Spearman entre o valor genotípico real e o valor genotípico predito pela metodologia REML/BLUP variou de 0,76 a 0,92 nas populações estudadas (4, 10, 30 e 100 genótipos). A população com quatro genótipos apresentou menor correlação (0,76 e 0,78) em comparação com as demais populações (Tabela 7).

Não houve grande alteração na média da Correlação de Spearman quando foram analisadas populações balanceadas e desbalanceadas (Tabela 7), indicando que o ranqueamento dos genótipos foi similar nas duas situações.

Tabela 7 - Média da Correlação de Spearman (CS) e Quadrado Médio do Erro (QME) entre valores genotípicos reais e preditos pela metodologia REML/BLUP de populações com diferentes tamanhos, balanceadas e desbalanceadas.

Tamanho da população	Balanceada		Desbalanceada	
	CS	QME	CS	QME
4	0,76	13.969,24	0,78	14.476,98
10	0,83	19.520,22	0,85	22.455,15
30	0,91	15.748,95	0,92	13.740,24
100	0,93	13.468,71	0,92	14.158,47

Fritsche Neto et al. (2010), em estudo com híbridos de milho, avaliando o método REML/BLUP na predição de valores genotípicos com dados balanceados e desbalanceados, concluíram que, em dados com até 20% de desbalanceamento esse método prediz com acurácia os valores genotípicos dessa cultura, mesmo na presença de interação genótipo x ambiente. Neste sentido, o método REML/BLUP permite trabalhar com certo grau de desbalanceamento dos dados, sem comprometer a eficiência do programa de

melhoramento. Esses resultados corroboram com os obtidos no presente estudo, que utilizou nível de desbalanceamento de 20%.

Observou-se que com o aumento do número de genótipos, ocorreu um aumento da Correlação de Spearman, sendo que as populações com 30 e 100 genótipos apresentaram maiores valores (Tabela 7). Nestas populações, a classificação dos genótipos foi mais próxima da classificação real.

Foi avaliada a proximidade entre o valor real e o predito pelo método REML/BLUP por meio do Quadrado Médio do Erro. Considerando todas as populações avaliadas, observou-se que Quadrado Médio do Erro foi maior para as populações desbalanceadas, exceto para população com 30 genótipos. Este resultado evidencia que em populações balanceadas a predição dos valores genotípicos é mais acurada (Tabela 7). Apesar da classificação dos genótipos ter apresentado pouca alteração, conforme as correlações de Spearman obtidas para as populações balanceadas e desbalanceadas (Tabela 7), os valores genotípicos preditos foram mais acurados para as populações balanceadas, apresentando, em geral menores valores para o Quadrado Médio do Erro (Tabela 7).

#### 4.4 AVALIAÇÃO GENÉTICA SOB DIFERENTES NÚMEROS DE CORTES

##### 4.4.1 Estimativas de parâmetros genéticos

Na população com quatro genótipos, as estimativas de herdabilidade obtidas pela metodologia REML/BLUP foram baixas, em comparação com a herdabilidade real, considerando todos os números de cortes testados. A estimativa que mais se aproximou do valor real para esta população foi a obtida com oito cortes, com porcentagem de erro de 20% (Tabela 8).

Os melhoristas de plantas geralmente usam o coeficiente de repetibilidade para inferir sobre a repetição da característica nos diferentes cortes realizados. Uma boa estimativa deste parâmetro depende de estimativas fidedignas dos componentes de variância. Segundo Basso et al. (2009), a repetibilidade é usada como medida do limite superior da herdabilidade e da confiabilidade da predição do valor genético, a partir de sucessivas medições no indivíduo.

As médias dos coeficientes de repetibilidade estimados, considerando a população de quatro genótipos nos diferentes cortes avaliados, variaram de 0,33 a 0,40, sendo necessários oito cortes para conseguir uma determinação de pelo menos 80%. O coeficiente de determinação em torno de 0,80 é utilizado por vários autores (CAVALCANTE et al., 2012; GARGNELUTTI FILHO, 2004; FERREIRA et al., 1999) para prever o valor genético ou genotípico dos genótipos para fins de seleção. Cruz et al. (2004) afirmam que valores baixos de estimativas de repetibilidade para determinado caráter exige um maior número de cortes ou medições para prever o valor genotípico mais próximo com maior precisão. Botrel et al. (2000) salientam que com coeficiente de repetibilidade alto é possível prever o valor genotípico dos indivíduos com menor número de cortes.

Tabela 8 - Média e desvio padrão (DP) de parâmetros genéticos estimados pela metodologia REML/BLUP de populações balanceadas com diferentes tamanhos em função do número de cortes

CORTES		TAMANHO DA POPULAÇÃO (indivíduos)							
		4		10		30		100	
		Média	DP	Média	DP	Média	DP	Média	DP
3	r	0,40	0,13	0,65	0,07	0,59	0,03	0,57	0,03
	$h^2_g$	0,09	0,11	0,21	0,06	0,21	0,06	0,23	0,03
	$R^2$	0,66	0,13	0,85	0,04	0,81	0,02	0,80	0,03
4	r	0,35	0,10	0,56	0,05	0,54	0,03	0,52	0,02
	$h^2_g$	0,16	0,11	0,23	0,05	0,24	0,05	0,26	0,03
	$R^2$	0,68	0,10	0,83	0,03	0,82	0,02	0,81	0,01
5	r	0,33	0,11	0,50	0,06	0,48	0,03	0,48	0,02
	$h^2_g$	0,17	0,11	0,24	0,06	0,24	0,05	0,27	0,03
	$R^2$	0,71	0,11	0,83	0,04	0,82	0,02	0,82	0,01
6	r	0,36	0,10	0,45	0,06	0,45	0,03	0,46	0,02
	$h^2_g$	0,22	0,10	0,24	0,06	0,25	0,05	0,29	0,02
	$R^2$	0,76	0,08	0,83	0,04	0,83	0,02	0,83	0,01
7	r	0,34	0,09	0,44	0,05	0,44	0,03	0,44	0,02
	$h^2_g$	0,20	0,08	0,27	0,05	0,28	0,05	0,29	0,02
	$R^2$	0,77	0,07	0,84	0,03	0,84	0,02	0,84	0,01
8	r	0,37	0,09	0,37	0,05	0,43	0,03	0,42	0,02
	$h^2_g$	0,24	0,08	0,24	0,05	0,29	0,04	0,30	0,02
	$R^2$	0,81	0,05	0,82	0,03	0,86	0,02	0,85	0,01

r: repetibilidade;  $h^2_g$ : herdabilidade individual;  $R^2$ : coeficiente de determinação.

Os desvios padrão das estimativas dos parâmetros genéticos para a população com quatro genótipos foi maior quando se utilizaram 3, 4, 5 e 6 cortes (Tabela 8), mostrando que houve uma maior dispersão dos valores de repetibilidade, herdabilidade e coeficiente de determinação em relação à média, quando analisou-se uma população pequena.

Para a população com 10 genótipos, as herdabilidades também foram subestimadas, sendo que ao se considerar sete cortes, obteve-se a estimativa de herdabilidade com menor porcentagem de erro (10%). As médias dos coeficientes de repetibilidade estimados para população com 10 genótipos nos diferentes cortes testados foram de mediana a alta magnitude (0,37 a 0,65). Com coeficiente de repetibilidade mais alto foi possível obter coeficiente de determinação maior que 0,80 a partir do terceiro corte (Tabela 8).

Os desvios padrão para população com 10 genótipos foram menores em comparação com os desvios padrão das populações com quatro genótipos, indicando que com maior número de observações há menor dispersão das estimativas de parâmetros genéticos em relação à média.

Considerando a característica produção de matéria seca de amendoim forrageiro pode-se inferir que, com 10 genótipos, o valor genotípico pode ser predito com 80% de confiança a partir de três cortes.

Nas populações com 30 e 100 genótipos, as herdabilidades estimadas foram se aproximando das herdabilidades reais à medida que se aumentou o número de cortes. Para população com 30 genótipos, quando considerados sete e oito cortes, as estimativas de herdabilidade apresentaram porcentagem de erro de 9,7% e 6,6%, respectivamente. Na população com 100 genótipos, a partir de avaliações considerando cinco cortes, as porcentagens de erro das estimativas de herdabilidade também foram baixas (6,90%, 0%, 0% e 3,4% para 5, 6, 7 e 8 cortes, respectivamente). Verificou-se, portanto, que em populações maiores há maior acurácia na estimação da herdabilidade a partir dos componentes de variância obtidos pelo REML.

Em relação aos coeficientes de repetibilidade estimados para população com 30 e 100 genótipos, estes foram de moderada a alta magnitude, variando de 0,42 a 0,59, considerando todos os cortes avaliados. Com esses valores de repetibilidade é possível obter confiança de 0,80 na predição do valor genético real a partir de três cortes (Tabela 8).

A utilização de um número de cortes relativamente pequeno para avaliação de plantas forrageiras, com confiança máxima, é de grande importância para os programas de melhoramento, pois possibilita menor gasto de tempo e mão de obra. No presente trabalho, observou-se que com o aumento da população, houve um aumento na confiabilidade da predição do valor real dos genótipos com reduzido número de cortes.

Estudos com dados experimentais são divergentes quanto ao número de cortes necessários para predizer o valor genético real do indivíduo com confiança de 0,80 para produção de matéria seca de forragem. Cavalcante et al. (2012), em estudo com 16 genótipos de capim-elefante concluíram que uma medição seria necessária para predizer o valor genético real com confiabilidade de 0,80. Este autor utilizou delineamento em blocos casualizados em esquema de parcelas subdivididas no tempo, com três blocos e seis cortes.

Daher et al. (2004) determinaram que nove avaliações seriam suficientes para seleção dos genótipos mais promissores de capim elefante. Estes autores utilizaram delineamento em blocos casualizados completos, em esquema de parcelas subdivididas no tempo, sendo 17 clones nas parcelas e 12 cortes nas subparcelas. Para Gargnelutti Filho et al. (2004), que avaliaram cinco genótipos de *Panicum maximum*, com 8 cortes a partir da média das repetições, em um modelo estatístico reduzido com dois efeitos (efeito de corte e genótipo), concluíram que seriam necessários 14 cortes para predizer o valor genético dos genótipos com 80% de confiança.

#### 4.4.2 Correlação de Spearman e quadrado médio do erro

As correlações de Spearman entre os valores genotípicos reais e preditos pelo método REML/BLUP para a população com quatro genótipos (Tabela 9) foram, em geral, menores em comparação com as populações de 10, 30 e 100 genótipos, considerando os diferentes números de cortes. Diferentemente das outras, esta população com quatro genótipos não apresentou aumento linear da correlação de Spearman conforme se aumentou o número de cortes, mostrando que em populações muito pequenas os resultados são inconsistentes (Tabela 9). Observou-se também que a classificação mais próxima da real foi quando se considerou a análise com oito cortes (Tabela 9). Este resultado evidencia que em populações menores é necessário um maior número de cortes para selecionar indivíduos geneticamente superiores com

maior confiança. Este resultado pode ser confirmado pela estimativa do coeficiente de determinação que só alcançou 80% de confiança quando considerados os oito cortes na análise (Tabela 8).

Tabela 9 - Correlação de Spearman (CS) e Quadrado Médio do Erro (QME) entre valores genotípicos reais e preditos pela metodologia REML/BLUP de diferentes populações balanceadas em função do número de cortes

Nº de cortes	POPULAÇÃO							
	4		10		30		100	
	CS	QME x 10 <sup>4</sup>	CS	QME x 10 <sup>4</sup>	CS	QME x 10 <sup>4</sup>	CS	QME x 10 <sup>4</sup>
3	0,74	5,65	0,62	5,51	0,85	2,78	0,81	3,21
4	0,64	3,76	0,71	4,20	0,87	2,28	0,85	2,57
5	0,62	2,93	0,76	3,28	0,89	1,98	0,88	2,14
6	0,60	2,02	0,76	2,97	0,91	1,65	0,90	1,77
7	0,74	2,03	0,80	2,48	0,92	1,44	0,92	1,52
8	0,76	1,30	0,83	1,95	0,91	1,57	0,93	1,34

Para as populações com 10 genótipos, observou-se que com o aumento do número de cortes a correlação de Spearman aumentou, chegando a 0,80 a partir do sétimo corte. Este resultado indica que a classificação dos genótipos se aproxima da classificação real com o aumento do número de cortes (Tabela 9).

As populações com 30 e 100 genótipos apresentaram correlação de Spearman similares considerando os diferentes números de cortes. Nestas populações, as correlações foram consideradas altas, sendo maior que 0,80 a partir da análise com três cortes. As correlações foram maiores que 0,90 ao se considerar seis cortes. Este resultado confirma que à medida que se aumenta o número de cortes, a classificação dos genótipos se aproxima da classificação real.

Souza Sobrinho et al. (2010), trabalhando com dados experimentais, avaliaram 118 genótipos de *Urochloa ruziziensis*, quanto ao número de cortes necessários para predizer o valor genético dos indivíduos. Os autores concluíram que seriam necessários de 8 a 10 avaliações para predizer o valor genotípico com 0,80 de confiança. Este resultado difere do apresentado no

presente estudo; porém, os autores utilizaram delineamento em blocos ao acaso, com duas repetições e sete cortes, sendo que na análise de variância não consideraram esquema de parcelas subdivididas no tempo.

Em relação ao Quadrado Médio do Erro, observou-se que para todas as populações analisadas (4, 10, 30 e 100 indivíduos) este diminuiu à medida que se aumentou o número de cortes, sendo menor com número de cortes igual a oito (Tabela 8), com exceção do quadrado médio do erro apresentado pela população com 30 genótipos, onde este se apresentou menor quando considerado sete cortes. Este resultado indica que quanto maior a população analisada, a metodologia REML/BLUP prediz com maior acurácia o valor genotípico dos indivíduos.

#### 4.5 SELEÇÃO COM BASE EM QUADRADOS MÍNIMOS E REML/BLUP

O ganho com a seleção é o principal objetivo do melhoramento genético. A resposta à seleção depende da coleta de dados fenotípicos de forma criteriosa e de estimativas de componentes de variância da população a ser melhorada. Sabe-se que procedimentos de análise com diferentes metodologias podem levar a diferenças na seleção dos indivíduos, tendo efeito direto no ganho genético.

Comparando-se a média do ganho genético real das populações de 4, 10, 30 e 100 genótipos com os dois níveis de desbalanceamento (0% e 20%), ranqueados pela metodologia de modelos mistos (REML/BLUP) e quadrados mínimos (teste de Tukey), observou-se maior ganho tanto para população balanceada quanto desbalanceada quando os genótipos foram classificados pelo REML/BLUP, exceto para população com quatro genótipos, onde o ganho genético com a seleção foi o mesmo comparando-se ambas as metodologias (Tabela 10 e Tabela 11).

Na população com quatro genótipos, um dos indivíduos simulados apresentou, de forma aleatória, valor genotípico muito superior aos demais, fato que possibilitou a identificação do melhor genótipo com maior facilidade pelas duas metodologias testadas.

Tabela 10 - Média ( $\bar{X}$ ) e desvio padrão (DP) do ganho genético real da produção de matéria seca (kg/ha) com a seleção em populações de diferentes tamanhos e dois níveis de desbalanceamento, utilizando o teste de Tukey

Tamanho da População	IS* (%)	Balanceada		IS(%)	Desbalanceada	
		$\bar{X}$	DP		$\bar{X}$	DP
4	25	415,00	0,00	25	415,00	0,00
10	28	421,61	34,23	31	400,12	55,07
30	20	426,60	40,23	23	416,06	38,74
100	15	465,58	21,94	15	465,58	15,97

\*Intensidade de Seleção

Tabela 11 - Média ( $\bar{X}$ ) e desvio padrão (DP) de ganho genético real da produção de matéria seca (kg/ha) com a seleção em populações de diferentes tamanhos e dois níveis de desbalanceamento, utilizando a metodologia REML/BLUP

População	Intensidade de seleção (%)	Balanceada		Desbalanceada	
		$\bar{X}$	DP	$\bar{X}$	DP
4	25	415,00	0,00	415,00	0,00
10	20	429,05	19,96	431,27	18,29
30	15	454,08	23,29	451,29	37,64
100	15	467,59	19,72	468,31	17,14

Observa-se também que ao considerar o teste de Tukey na seleção dos genótipos, as populações desbalanceadas com 10 e 30 indivíduos apresentaram ganho genético menor quando comparadas às populações balanceadas (Tabela 10). Por outro lado, os ganhos verificados ao se utilizar o método REML/BLUP foram bastante similares nas populações com 0 e 20% de desbalanceamento. Esta é uma vantagem do REML/BLUP em relação aos testes de médias, sendo que este método permite trabalhar com dados desbalanceados (RESENDE, 2007).

Resende e Duarte (2007) evidenciam que o uso de modelos fixos para avaliação de tratamentos genéticos somente é adequado com número de tratamentos inferior ou igual a cinco, sendo que o modelo com tratamento aleatório pode ser usado de forma generalizada. O mesmo autor salienta que o uso do modelo aleatório reduz o risco de erros na inferência do valor genotípico dos indivíduos. No presente trabalho, considerando a população com quatro genótipos, o uso do modelo fixo (metodologia de quadrados mínimos) e misto

(REML/BLUP) não alterou os ganhos com a seleção, sendo considerados iguais no caso de populações balanceadas ou com 20% de desbalanceamento.

Nas populações balanceadas com 10 genótipos, o uso do método REML/BLUP proporcionou um ganho genético 1,76% maior (Tabela 10 e Tabela 11) quando comparado com a seleção dos genótipos realizada com base no teste de médias de Tukey. Na população desbalanceada, a seleção baseada no método REML/BLUP proporcionou ganho genético 7,78% maior (Tabela 10 e Tabela 11) em relação ao ganho obtido quando se utilizou teste de médias.

Mendes et al. (2012), em estudo comparando os métodos REML/BLUP e quadrados mínimos na cultura do feijoeiro, em experimento em blocos casualizados completos, confirmam que os dois métodos testados levam a classificações distintas quando considerado um grande número de genótipos. Estes autores avaliaram o método de quadrados mínimos (LS) com diferentes estratégias de seleção (seleção massal, massal estratificada e seleção entre e dentro de progênie) com relação à classificação genotípica realizada pelo método BLUP. Os resultados deste trabalho mostram que ao considerar os primeiros 10 indivíduos classificados por LS e BLUP houve 100% de porcentagem de coincidência entre a classificação dos dois métodos. Porém, considerando os 100 melhores indivíduos selecionados por seleção massal e massal estratificada, a porcentagem de coincidência em relação ao BLUP não foi superior a 40%.

Vale ressaltar que no presente trabalho, quando se utilizou teste de médias, a intensidade de seleção foi maior, pois a mesma foi determinada pelo resultado do teste de Tukey. Este teste não foi capaz de diferenciar os genótipos, quando na realidade eram distintos, ocasionando o erro tipo II, que é definido como a probabilidade de aceitar uma hipótese nula dado que ela é falsa. Segundo Hayter (1984), o teste de Tukey é extremamente conservador, que embora controle bem o erro tipo I, permite o aparecimento do erro tipo II. O mesmo autor afirma que em dados desbalanceados, este teste leva a resultados aproximados, o que pode ocasionar a escolha de genótipos não superiores geneticamente.

A metodologia REML/BLUP prediz os valores genéticos, ponderando-os concomitantemente com a estimação dos efeitos fixos (RESENDE, 2007). Este fato proporciona uma classificação hierárquica dos genótipos permitindo ao

melhorista flexibilidade de decisão em relação à intensidade de seleção praticada. Por ser um teste rigoroso, o teste de Tukey conduz a uma maior incidência de erro tipo II, dificultado a distinção entre genótipos e, conseqüentemente, a decisão sobre o material genético a ser selecionado.

Carvalho et al. (2008) observaram resultados diferentes dos encontrados neste trabalho em dados experimentais de genótipos de soja, quantificando a resposta com a seleção pelos métodos REML/BLUP e quadrados mínimos. Estes autores mostraram que não houve diferença significativa entre o ganho obtido pelas duas metodologias. Os autores consideraram que as pequenas diferenças entre ganhos não eram relevantes para o melhoramento desta cultura. Porém, no estudo de Carvalho et al. (2008) foram utilizados dados experimentais, não sendo possível conhecer o verdadeiro valor genético dos genótipos.

A superioridade, em geral, observada do método REML/BLUP em relação à metodologia de quadrados mínimos para o ganho com a seleção de genótipos superiores, é de grande importância para o melhoramento de plantas forrageiras tropicais. Considerando que o tempo para obtenção de novas cultivares forrageiras é longo, a utilização do método REML/BLUP pode trazer grandes vantagens no decorrer dos ciclos de seleção.

Duarte e Vencovsky (2001), em abordagem teórica sobre o uso do método REML/BLUP e quadrados mínimos no ordenamento de materiais genéticos, ressaltam que esses modelos levam a diferentes ranqueamentos no caso de dados desbalanceados se os tratamentos diferirem muito em número de repetições, ou quando o experimento for em blocos incompletos. Estes autores afirmam também que quando a variância genotípica é alta, o ordenamento dos genótipos pelas duas metodologias é idêntico ou muito similar.

Para população com 30 genótipos, o uso do método REML/BLUP proporcionou um ganho superior de 6,44% na população balanceada e 8,44% na população desbalanceada em relação ao teste de Tukey (Tabela 10 e Tabela 11). Observou-se que com o aumento do número de genótipos há uma maior vantagem na utilização da metodologia REML/BLUP, pois o ganho com a seleção é maior nas populações balanceadas e desbalanceadas.

Estudos mostram que com o aumento do número de tratamentos, o poder de diferenciação entre os mesmos, utilizando-se o teste de Tukey,

diminui. Este fato é potencializado pelo aumento do coeficiente de variação (BORGES; FERREIRA, 2003; GIRARD et al., 2009). Além disso, Resende e Duarte (2007) afirmam que testes baseados no método de quadrados mínimos, considerando efeito de tratamento como fixo, só devem ser usados quando considerados até cinco genótipos.

Apesar de se verificar nos estudos de melhoramento de plantas forrageiras, altos coeficientes de variação e perdas de plantas ou parcelas, este teste é de uso comum entre os pesquisadores, para um grande número genótipos. Essa situação pode ser observada nos estudos de Ferreira et al. (1999,) avaliando 42 cultivares de alfafa; Lopes et. al. (2010), avaliando 22 genótipos de palma forrageira quanto à resistência das cultivares a pragas; e Monteiro et. al. (2004), avaliando o desempenho 51 genótipos de sorgo forrageiro. No presente estudo, observou-se que o uso do teste de Tukey na avaliação de experimentos com número de genótipos maior ou igual a 10, proporcionou menor ganho com a seleção em populações balanceadas e desbalanceadas, em comparação com o método REML/BLUP.

Na população balanceada com 100 genótipos, a seleção feita pela média fenotípica proporcionou mesmo ganho genético quando consideradas as populações balanceada e desbalanceada (Tabela 10). Ao comparar as metodologias REML/BLUP e quadrados mínimos, observou-se que ambas proporcionaram ganhos similares nas populações balanceada e desbalanceada, com pequena vantagem para o método REML/BLUP, que proporcionou aproximadamente 1% a mais de ganho genético real (Tabela 10 e Tabela 11).

Em relação aos desvios padrão das médias do ganho genético real, observou-se que a população com quatro genótipos apresentou desvio padrão nulo, já que para essa população a metodologia REML/BLUP e o teste de Tukey obtiveram o mesmo ganho genético nas 10 repetições realizadas.

Para as populações balanceadas e desbalanceadas com 10, 30 e 100 genótipos, os desvios padrão em relação à média de ganho foram maiores quando se utilizou o teste de Tukey para seleção dos genótipos fenotipicamente superiores, com exceção da população desbalanceada com 100 genótipos, onde o método REML/BLUP apresentou desvio padrão ligeiramente superior (Tabela 10 e Tabela 11). Os desvios padrão mais elevados mostram que há uma maior dispersão das médias fenotípicas dos

genótipos selecionados pela metodologia de quadrados mínimos em relação à média total.

O método REML/BLUP pode ser mais eficiente devido à seleção ser baseada em valores genotípicos, sendo retirados os efeitos ambientais presentes nos valores fenotípicos observados. Vários estudos com dados experimentais confirmam a superioridade deste método quanto à predição de valores genéticos mais acurados, dentre eles pode-se citar Resende et al. (1996) e Resende e Duarte (2007).

Estudos mostram que o fato de o pesquisador utilizar modelos diferentes nas análises estatísticas, pode levar a resultados diferentes no ordenamento dos genótipos. Scarpinati et al. (2009), em estudo sobre a influência do modelo de análise estatística e da forma das parcelas experimentais na seleção de clones de *Eucalyptus* spp., testaram as metodologias de quadrados mínimos e REML/BLUP quanto ao ordenamento dos clones e confirmaram que, ao utilizar modelo com clone fixo e bloco aleatório e clone aleatório e bloco fixo, em experimento em blocos com uma planta por parcela, houve pequena diferença quanto ao ordenamento dos clones. Estes resultados corroboram com apresentados no presente estudo, exceto para população com quatro genótipos, onde o ranqueamento para as duas metodologias foi igual.

## 5 CONCLUSÕES

As metodologias REML e quadrados mínimos geram estimativas similares de parâmetros genéticos, em populações balanceadas e desbalanceadas de amendoim forrageiro, com diferentes números de indivíduos.

O uso da metodologia REML/BLUP conduz a maiores ganhos de seleção quando comparada à metodologia de quadrados mínimos com uso do teste de Tukey, exceto para populações pequenas, cujo ganho é o mesmo para ambos os métodos de análise.

A classificação dos genótipos superiores com base em seus valores genotípicos preditos é prejudicada quando a população é pequena.

A acurácia da predição dos valores genotípicos é, em geral, maior em populações balanceadas. Porém, o nível de desbalanceamento de 20% não interfere na seleção dos genótipos superiores para as condições simuladas neste estudo.

Para populações pequenas é necessário um maior número de cortes para predição dos valores genotípicos dos indivíduos. Contudo, para populações a partir de 10 genótipos, três cortes são suficientes para predizer o valor genotípico com coeficiente de determinação de 80%.

## REFERÊNCIAS

ANDRADE, V. T. **Implicações da modelagem estatística na seleção de progênies de cafeeiro**. Lavras, 2013. Dissertação (Mestrado), Universidade Federal de Lavras, 2013.

ARAÚJO S. A. C; DEMINICIS B.B; CAMPOS P.R.S.S. Melhoramento genético de plantas forrageiras tropicais no Brasil. **Archivos de Zootecnia**, 57:61-76. 2008.

ASSIS, G. M. L. de. Melhoramento de leguminosas forrageiras tropicais. In: SIMPÓSIO SOBRE MANEJO ESTRATÉGICO DA PASTAGEM, 5.; SIMPÓSIO INTERNACIONAL SOBRE PRODUÇÃO ANIMAL EM PASTEJO, 3., Viçosa. **Anais...** Viçosa, MG: UFV, 2010. P. 213-249.

ASSIS, G. M. L. de; CAMPOS, T. de; BIANCHINI, P. C.; MATOS, L. R. A. de; Banco de germoplasma de amendoim forrageiro: conservação e utilização. In: CONGRESSO BRASILEIRO DE RECURSOS GENÉTICOS, 2, Belém, PA. **Anais...** Brasília, DF: Sociedade Brasileira de Recursos Genéticos, 2012. CD-ROM.

ASSIS, G. M. L. de; VALENTIM, J. F. Programa de melhoramento genético do amendoim forrageiro: avaliação agrônômica de acessos no Acre. **Amazônia: Ciência e Desenvolvimento**, v. 4, n. 8, 2009.

ASSIS, G. M. L. de; VALENTIM, J. F. Forage peanut breeding program in Brazil. In: SIMPÓSIO INTERNACIONAL SOBRE MELHORAMENTO DE FORRAGEIRAS, 2, Campo Grande, MS. **Anais...** Campo Grande: Embrapa Gado de Corte, 2009. CD-ROM.

ASSIS, G. M. L. de; VALENTIM, J. F.; CARNEIRO JÚNIOR, J. M.; AZEVEDO, J. M. A de; FERREIRA, A. S. Seleção de genótipos de amendoim forrageiro para cobertura do solo e produção de biomassa aérea no período de estabelecimento utilizando-se metodologia de modelos mistos. **Revista Brasileira de Zootecnia**, v. 37, n. 11, p. 1905-1911, 2008.

ASSIS, G. M. L. de; VALLS, J. F. M.; CARVALHO, M. A.; VALENTIM, J. F.; ANDRADE, C. M. S. de. **Descritores morfológicos para condução de ensaios de distinguibilidade, homogeneidade e estabilidade em *Arachis pintoi* Krapov. & W. C. Greg.** Rio Branco, AC: Embrapa Acre, 2010, 25 p. (Documentos, 117).

AZEVEDO, J. M. A. de; ASSIS, G. M. L. de; SAGGIN JUNIOR, O. J.; AZEVEDO, H. S. F. da S. Riqueza e frequência de espécies de fungos micorrízicos arbusculares em genótipos de amendoim forrageiro no Acre, Norte do Brasil. **Acta Amazônica**, v. 44, n. 2, p. 157-168, 2014.

BARBIN, D. **Componentes de variância**. Piracicaba: ESALQ, USP, 1993. 108p.

BARCELOS, A. de O.; VILELA, L.; LUPINACCI, A.V. Produção animal a pasto: desafios e oportunidades. In: ENCONTRO NACIONAL DO BOI VERDE: A

- PECUÁRIA SUSTENTÁVEL, 3., 2001. Uberlândia. **Anais...** Sindicato Rural de Uberlândia, 2001. p. 20-64.
- BASSO, K. C.; RESENDE, R. M. S.; VALLE, C. B.; GONÇALVES, M. C.; LEMPP, B. Avaliação de acessos de *Brachiaria brizantha* Stapf e estimativas de parâmetros genéticos para caracteres agronômicos. **Acta Scientiarum. Agronomy**, Maringá, v. 31, n. 1, p. 17-22, 2009.
- BORGES, L. C.; FERREIRA, D. F. Poder e taxas de erro tipo I dos testes de Scott-knott, Tukey e Student-Newman-Keuls sob distribuições normais e não normais dos resíduos. **Revista de Matemática e Estatística**, v. 21, n. 1, p. 67-83, 2003.
- BOTREL, M. A.; FERREIRA, R. P.; CRUZ, C. D.; PEREIRA, A. V.; VIANA, M. C. M.; ROCHA, R.; MIRANDA, M. Estimativas de coeficientes de repetibilidade para produção de matéria seca em cultivares de alfafa, sob diferentes ambientes. **Revista Ceres**, v. XLVII, n. 274, p.651-663, 2000.
- BUENO FILHO, J. S. S. **Modelos mistos na predição de valores genéticos aditivos em teste de progênies florestais**. Tese (Doutorado em Genética e melhoramento de Planta) – Universidade de São Paulo, Escola Superior de Agricultura Luiz Queiroz, Piracicaba, 1997. 118 f.
- CAMARINHA FILHO, J. A. **Modelos lineares mistos: estruturas de matrizes de variância e covariância e seleção de modelos**. Tese (Doutorado em Estatística e Experimentação Agronômica) – Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2002, 96 p.
- CARNEIRO JÚNIOR, J. M. **Avaliação da heterogeneidade de variâncias utilizando dados simulados**. Tese (Doutorado em genética e melhoramento), Universidade Federal de Viçosa, Viçosa, 2005.
- CARNEIRO JÚNIOR, J. M.; EUCLYDES, R. F.; LOPES, P. S.; TORRES, R. A. Avaliação de métodos de estimação de componentes de variância utilizando dados simulados. **Revista Brasileira de Zootecnia**, v. 33, n. 2, p. 328-336, 2004.
- CARVALHO, A. D. F. de; FRITSCH NETO, R.; GERALDI, I. O. Estimation and prediction of parameters and breeding values in soybean using REML/BLUP and Least Squares. **Crop Breeding and Applied Biotechnology**, v.8, p. 219-224, 2008.
- CAVALCANTE, M.; LIRA, M. A.; SANTOS, M. V. F.; PITA, E. B. A. F.; FERREIRA, R. L. C.; TABOSA, J. N. Coeficiente de repetibilidade e parâmetros genéticos em capim-elefante. **Pesquisa Agropecuária Brasileira**, v. 47, n. 4, p. 569-575, abr. 2012.
- CECON, P. R.; SILVA, F. F.; FERREIRA, A.; FERRÃO, R. G.; CARNEIRO, A. P. S.; DETMANN, E.; FERIA, P. N.; MORAIS, T. S. S. Análise de medidas repetidas na avaliação de clones de café 'Conilon'. **Pesquisa Agropecuária Brasileira**, v. 43, n. 9, p. 1171-1176, set. 2008.

COELHO, A. M. **Simulação de dados visando a estimação de componentes de variância e coeficientes de herdabilidade**. Piracicaba, 2005. Dissertação (Mestrado), Escola Superior de Agricultura Luiz Queiroz, 2005.

COELHO, A. M.; BARBIN, D. Simulação de dados visando a estimação de componentes de variância e coeficientes de herdabilidade. **Revista de Matemática e Estatística**, v. 24, n. 2, p. 103-120, 2006.

CRUZ, C. D. **A informática no melhoramento genético**. In: NASS, L. L. et al. (Eds). Recursos Genéticos e Melhoramento de Plantas. Rondonópolis: Fundação-MT, 2001. P. 1086-1118.

CRUZ, C. D. **Programa genes versão Windows**: aplicativo computacional em genética e estatística. Viçosa: UFV, 2001. 648 p.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. 3. ed. Viçosa: UFV, 2004. 480 p.

DAHER, R. F.; MALDONADO, H.; PEREIRA, A. V.; AMARAL JUNIOR, A. T.; PEREIRA, M. G.; FERREIRA, C. F.; RAMOS, S. R. R.; TARDIN, F. D.; SILVA, M. P. Estimativas de parâmetros genéticos e de coeficientes de repetibilidade de caracteres forrageiros em clones de capim elefante (*Pennisetum purpureum Schum*). **Acta Scientiarum. Agronomy**, v. 26, n. 4, p. 483-490, 2004.

DIAS FILHO, M. B. **Diagnóstico das pastagens no brasil**: pastagens nativas, cultivadas e degradadas. Disponível em: <<http://boiapasto.com.br/2013/08/diagnostico-da-situacao-das-pastagens-no-brasil-pastagens-nativas-cultivadas-e-degradadas/>>. Acesso em: 09 Março 2014.

DUARTE, J. B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal**. 2000. 293 p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz Queiroz, USP, Piracicaba, 2000.

DUARTE, J. B.; VENCOVSKY, R. Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. **Scientia Agricola**, v.58, n.1, p.109-117, 2001.

DUARTE, J. B.; VENCOVSKY, R.; DIAS, C. T. S. Estimadores de componentes de variância em delineamento de blocos aumentados com tratamentos novos de uma ou mais populações. **Pesquisa Agropecuária Brasileira**, v. 36, n. 9, p. 1155 – 1167, 2001.

EISENHART, C. The assumptions underlying the analysis of variance. **Biometrics**, Washington, v. 3, n. 1, p. 1-21, 1947.

ESTAGHVIROU, S. B. O.; OGUTU, J. O.; SCHULZ-STREECK, T. KNAAK, C.; OUZUNOVA, M.; GORDILLO, A.; PIEPHO, H-P. Evaluation of approaches for estimating the accuracy of genomic prediction in plant breeding. **BMC Genomics**, v. 14, n. 860, 2013.

EUCLIDES, V. P. B.; EUCLIDES FILHO, K.; COSTA, F. P.; FIGUEIREDO, G. R. Desempenho de novilhos F1s angus-nelore em pastagens de *Brachiaria decumbens* submetidos a diferentes regimes alimentares. **Revista Brasileira de Zootecnia**, Viçosa, v. 30, n. 2, p. 470-481, 2001.

FARIA, V. R. **Melhor predição linear imparcial (BLUP) no melhoramento vegetal: seleção entre famílias de meios-irmãos**. 2008 57 f. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa, 2008.

FARIAS NETO, J. T. de; RESENDE, M. D. V. de. Aplicação da metodologia de modelos mistos (REML/BLUP) na estimação de componentes de variância e predição de valores genéticos em pupunheira (*Bactris gasipaes*). **Revista Brasileira de Fruticultura**, v. 23, n. 2, p. 320-324, 2001.

FERNANDEZ, G.C.J. Repeated measure analysis of line-source sprinkler experiments. **Hort Science**, v.26, n.4, p.339-342, 1991.

FERREIRA, A. L.; MAURÍCIO, R. M.; FERNANDES, F. D.; CARVALHO, M. A.; RAMOS, A. K. B.; GUIMARÃES JUNIOR, R. Ranking contrasting genotypes of forage peanut based on nutritive value and fermentation kinetics. *Animal Feed Science and Technology*, n. 175, p. 16-23, 2012.

FERREIRA, A.; CRUZ, C. D.; VASCONCELOS, E. S. de; NASCIMENTO, M.; RIBEIRO, M. F.; SILVA, M. F. da. Utilização de bootstrap não-paramétrico para avaliação de correlações fenotípicas, genotípicas e ambientais. **Acta Scientiarum. Agronomy**, Maringá, v. 30, p. 657-663, 2008.

FERREIRA, D. F. **Eficiência de métodos de mapeamento de locos quantitativos (QTLs) e da seleção assistida por marcadores moleculares através de simulação**. 1995 210 p. Tese (Doutorado) – Escola Superior de Agricultura Luiz Queiroz, Piracicaba, 1995.

FERREIRA, D. F. Uso de simulação no melhoramento. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S.; VALADARES-INGLIS, M. C. (Ed.). **Recursos Genéticos e Melhoramento**. Rondonópolis: Fundação MT, 2001. Cap. 35, p.1120-1139.

FERREIRA, R. de P.; BOTREL, M. de A.; PEREIRA, A. V.; CRUZ, C. D. Avaliação de cultivares de alfafa e estimativas de repetibilidade de características forrageiros. **Pesquisa Agropecuária Brasileira**, v. 34, n. 6, p. 995-1002, 1999.

FIGUEIREDO, J. U. de. **Estimação de parâmetros genéticos e fenotípicos em progênies de *Brachiaria humidicola***. 2011 75 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Lavras, Lavras, 2011.

FREITAS FILHO, P. J. de. **Introdução à Modelagem e Simulação de Sistemas - com Aplicações em Arena**. 2 ed. Florianópolis: Visual Books, 2008. v. 2000. 380p

FREITAS, A. R.; FERREIRA, R. P.; MOREIRA, A. Análise de dados de medidas repetidas por meio do modelo linear geral e do modelo misto. **Revista de Ciências Agrárias**, v. 54, n. 3, p. 214-224, set./dez. 2011.

FREITAS, E. G. de; BARBIN, D.; BARBOSA, G. V. de S.; CARNEIRO, M. S.; BASSINELLO, A. I. Modelo univariado aplicado a dados longitudinais de cana-de-açúcar. **Revista Brasileira de Biometria**, v. 26, n. 2, p. 93-106, 2008.

FRITSCH NETO, R.; GONÇALVES, M. C.; VENCOVSKY, R., SOUZA JUNIOR, C. L. Prediction of genotypic values of maize hybrids in unbalanced experiments. **Crop Breeding and Applied Biotechnology**, 10: 32-39, 2010.

FURTINI, I. V. **Estratégias de seleção em testes de progênie visando a obtenção de clones de eucalipto para a indústria de celulose**. Tese (Doutorado em Genética e melhoramento de plantas), Universidade Federal de Lavras. Lavras: UFLA, 2011. 111 p.

GARGUINELUTTI FILHO A.; CASTILHOS, Z. M. de S.; STORCK, L.; SAVIAN, J. F. Análise de repetibilidade de caracteres forrageiros de genótipos de *Panicum máximo*, avaliados com e sem restrição solar. **Ciência Rural**, v. 34, n. 3, p. 723-729, 2004.

GIRARDI, L. H.; CARGUINELUTTI FILHO, A.; STORCK, L.; Erro tipo I e poder de cinco testes de comparação múltipla de médias. **Revista Brasileira de Biometria**, v. 27, n. 1, p. 23-36, 2009.

HAYTER, A. J. A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. **The Annals of Statistics**, v. 12, n. 1, p. 61-75, 1984.

HENDERSON, C. R. Beste Linear unbiased estimation and prediction under a selection model. **Biometrics**, Alexandria, v. 31, n. 2, p. 423-447, June, 1975.

HENDERSON, C.R. Estimation of variance and variance components. **Biometrics**, v.9, p.226-252, 1953.

HERNDERSON, C. R. Estimation of changes in herd environment. **Journal of Dairy Science**, v. 32, p. 709, 1949.

HUYNH, H.; FELDT, L.S. Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. **Journal of the American Statistical Association**, v.65, n.332, p. 1582-1589, Dez 1970.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção da pecuária municipal**. Rio de Janeiro, v. 40, p. 1-71, 2012.

KELTON, W. D.; SADOWSKI, R. P. E.; STURROCK, D. T. **Simulation with Arena**. 4. ed. New York: McGraw-Hill, 2007.

KERRIDGE, P.C. Perspectivas futuras para el uso y estudio de especies forrajeras de *Arachis*. In: KERRIDGE, P.C. (Ed.). **Biología y agronomía de**

**especies forrajeras de *Arachis***. Cali: Centro Internacional de Agricultura Tropical, 1995. p.215-223.

LASCANO, C. E. Nutritive value and animal production of forage *Arachis*. In: KERRIDGE, P. C.; HARDY, B. (Ed.). **Biology and agronomy of forage *Arachis***. Cali, Colômbia: Centro Internacional de Agricultura Tropical (CIAT), 1994. p. 109-121.

LÁZARO, C. C. M.; RAMOS, L. M.; RODRIGUES, T. de J. D.; GARGNELITTI FILHO, A.; MINARI, D. P.; PATERNIANI, M. L. S. Inferência do sombreamento no desempenho de genótipos de *Stylosanthes guianensis*. **Científica**, Jaboticabal, v. 37, n. 1, p. 1- 8, 2009.

LÉDO, F. J. da S.; PEREIRA, A. V.; BOTREL, M. de A.; SOUZA SOBRINHO, F. de.; OLIVEIRA, J. S. e; XAVIER, D. F.; HEINEMANN, A. B.; FERREIRA, R. de P. Avaliação de cultivares de alfafa na zona da mata de Minas Gerais. **Ciência e Agrotecnologia**, Lavras, v. 28, n. 5, p. 1151-1159, set./out. 2004.

LOPES, E. B.; BRITO, C. H. de; ALBUQUERQUE, I. C. de; BATISTA, J. de L. Seleção de genótipos de palma forrageira (*Opuntia spp.*) e (*Nopalea spp.*) resistentes a cochonilha-do-carmim (*Dactylopius opuntiae* Cockerell, 1929) na paraíba Brasil. **Engenharia Ambiental**, v. 7, n. 1, p. 204-215, 2010.

MACEDO, M. C. M. Pastagens no ecossistema cerrados: evolução das pesquisas para o desenvolvimento sustentável. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 42., 2005, Goiânia. **Anais...** Goiânia: SBZ, UFG, 2005. p. 56-84.

MACHADO, A. N.; SIEWERDT, L.; ZONTA, É. P.; VAHL, L. C.; COELHO, R. W.; FERREIRA, O. G. L.; AFFONSO, A. B. Rendimento do amendoim-forrageiro estabelecido sob diferentes arranjos populacionais de plantas em planossolo. **Ciência Animal Brasileira**, v. 6, n. 3, p. 151-162, 2005.

MALHEIROS, E. B. Precisão de testes F univariados usados em experimentos com medidas repetidas no tempo, quando a condição de esfericidade da matriz de covariâncias não é verificada. **Revista de Matemática e Estatística**, v. 22, n. 2, p. 23-29, 2004.

MALHEIROS, E.B. Precisão da análise de dados longitudinais, com diferentes estruturas para a matriz de variâncias e covariâncias, quando se utiliza o esquema em parcelas subdivididas. **Revista de Matemática e Estatística**, v.17, p.263-273, 1999.

MARCELINO, S. D. do R.; IEMMA, A. F. Métodos de estimação de componentes de variância em modelos mistos desbalanceados. **Scientia Agricola**, v.57, n.4, p.643-652, 2000.

MARIGUELE, K. H.; RESENDE, M. D. V. de; VIANA, J.M. S.; SILVA, F. F. e; SILVA, P. S. L. de; KNOP, F. C. Métodos de análise de dados longitudinais para o melhoramento genético da pinha. **Pesquisa Agropecuária Brasileira**, v. 46, n. 12, p. 1657-1664, 2011.

MARTINEZ, D. T. **Avaliação genética sob heterogeneidade de variância residual dentro de tratamentos**. 2010. 73 p. Tese (Doutorado em Ciências Florestais) – Universidade Federal do Paraná, Curitiba, 2010.

MARTUSCELLO J. A.; JANK L.; FONSECA D. M.; RESENDE R. M. S. Repetibilidade de caracteres agronômicos em *panicum maximum* jacq. **Revista Brasileira de Zootecnia**, 36:6, 2007

MAUCHLY, J. W. Significance test for sphericity of a normal n-variate distribution. **Annals of Mathematical Statistics**, v. 11, p. 204-209, 1940. I

MENDES, M. P.; RAMALHO, M. A. P.; ABREU, A. F. B. Strategies in identifying individuals in segregant population of common bean and implications of genotype x environment interaction in the success of selection. **Journal Genetics and Molecular Research**, v. 11, n. 2, p. 872-880, 2012.

MILES, J. W. Achievements and perspectives in the breeding of tropical grasses and legumes. In: INTERNATIONAL GRASSLAND CONGRESS, 19, 2001. São Pedro. **Proceedings...** Fundação de Estudos Agrários Luiz Queiroz, 2001. p. 509-515.

MIRANDA, C. H. B.; VIEIRA, A.; CADISCH, G. Determinação da fixação biológica de nitrogênio no amendoim forrageiro (*Arachis* spp.) por intermédio da abundância natural de <sup>15</sup>N. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 32, n. 6, p. 1859-1865, nov./dez. 2003.

MIRANDA, E. M. de.; SAGGIN JÚNIOR, O. J.; SILVA, E. M. R. da. **Amendoim forrageiro: importância, usos e manejo**. Seropédica: Embrapa Agrobiologia, 2008. 92 p. (Documentos, 259).

MONTEIRO, M. C. D.; ANUNCIÇÃO FILHO, C. J. da; TABOSA, J. N.; OLIVEIRA, F. J. de; REIS, O. V. dos; BASTOS, G. Q. Avaliação do desempenho de sorgo forrageiro para o semi-árido de Pernambuco. **Revista Brasileira de Milho e Sorgo**, v. 3, n. 1, p.52-62, 2004.

MORGAN, B. J. T. **Elements of simulation**. London: Chapman e Hall, 1995. 351 p.

ORTIZ, J. de O.; FELGUEIRAS, C. A.; DRUCK, S.; MONTEIRO, A. M. V. Modelagem de fertilidade do solo por simulação estocástica com tratamento de incertezas. **Pesquisa Agropecuária Brasileira**, v. 39, n. 4, p. 379-389, abr. 2004.

PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when blocks sizes are unequal. **Biometrika**, Cambridge, v. 58, n. 2, p. 545-554, Dec. 1971.

PEREIRA, A. V.; VALLE, C. B.; FERREIRA, R. P.; MILES, J. W. Melhoramento de forrageiras tropicais. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S.; VALADARES-INGLIS, M. C. (Ed.). **Recursos Genéticos e Melhoramento**. Rondonópolis: Fundação MT, 2001. Cap. 18, p.549-601.

PEREIRA, A.V. Avanços no melhoramento genético de gramíneas forrageiras tropicais. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 39, 2002, Recife. **Anais...** Recife: SBZ, 2002. p.19-41.

PEREZ, N. B.; PIZARRO, E. A. **Potencial forrajero del género *Arachis* en el trópico americano**. IX Seminario de Pastos y Forrajes. 2005. Disponível em: <[http://www.avpa.ula.ve/eventos/ix\\_seminario\\_pastosyforraje/Conferencias/C2-Nay losBastiani.pdf](http://www.avpa.ula.ve/eventos/ix_seminario_pastosyforraje/Conferencias/C2-Nay%20losBastiani.pdf)>. Acesso em: 24 abril 2014.

PIEPHO, H. P.; MOHRING, J. Computing heritability and selection response from unbalanced plant breeding trials. **Genetics**, n. 177, p. 1881- 1888, 2007.

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

RESENDE, M. D. V. de. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília, DF: Embrapa Informação Tecnológica, 2002. 975 p.

RESENDE, M. D. V. de. **Inferência bayesiana e simulação estocástica (amostragem de gibbs) na estimação de componentes de variância e valores genéticos em plantas perenes**. Colombo: Embrapa Florestas, 68 p. 2000. (Documentos 46).

RESENDE, M. D. V. de. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Brasília, DF: Embrapa Informação Tecnológica, 2007. 561 p.

RESENDE, M. D. V. de. **Selegen-Reml/Blup: Sistema Estatístico e Seleção Genética Computadorizada via Modelos Lineares Mistos**. Colombo: Embrapa Florestas, 2007a. 361 p.

RESENDE, M. D. V. de; BARBOSA, M. H.P. Selection via simulated individual BLUP based on family genotypic effects in sugarcane. **Pesquisa Agropecuária Brasileira**. v. 41, n. 3, p. 421-429, 2006.

RESENDE, M. D. V. de; PRATES, D. F.; JESUS, A.; YAMADA, C. K. Melhor predição linear não viciada (BLUP) de valores genéticos no melhoramento de *Pinus*. **Boletim de Pesquisa Florestal**, Colombo, n. 32/33, p. 3-22, 1996a.

RESENDE, M. D. V. de; PRATES, D. F.; YAMADA, C. K.; JESUS, A. Estimação de componentes de variância e predição de valores genéticos pelo método da máxima verossimilhança restrita (REML) e melhor predição linear não viciada (BLUP) em *Pinus*. **Boletim de Pesquisa Florestal**, Colombo, n. 32/33, p. 23-42, 1996.

RESENDE, R. M. S.; VALLE, C. B.; JANK, L.; **Melhoramento de forrageiras tropicais**. Campo Grande, MS: Embrapa Gado de Corte, 2008. 293 p.

ROCHA, M. G. B. et al. Avaliação genética de progênies de meio-irmãos de *Eucalyptus grandis* por meio dos procedimentos REML/BLUP e da ANOVA . **Scientia Florestalis**, Piracicaba, n. 71, p. 99-107, 2006a.

ROCHA, M. G. B. et al. Avaliação genética de progênies de meio-irmãos de *Eucalyptus urophylla* utilizando os procedimentos REML/BLUP e E(QM) . **Ciência Florestal**, v. 16, n. 4, p. 369-379, 2006b.

ROCHA, R. B.; RAMALHO, A. R.; TEIXEIRA, A. L.; LAVIOLA, B. G.; SILVA, F. C. G.; MILITÃO, J. S. L. T. Eficiência da seleção para incremento do teor de óleo do pinhão-manso. **Pesquisa Agropecuária Brasileira**, v. 47, n. 1, p. 44-50, 2012.

ROSADO, A. M. et al. Ganhos genéticos preditos por diferentes métodos de seleção em progênies de *Eucalyptus urophylla*. **Pesquisa Agropecuária Brasileira**, v. 44, n.12, p. 1653-1659. 2009.

ROSADO, A. M. **Seleção entre e dentro de famílias e baseada nos valores genéticos obtidos pelo índice combinado e blup em eucalipto**. 2003. Tese (Doutorado) - Universidade Federal de Viçosa, Viçosa,2003.

SANTOS, E. C. dos. **Características agronômicas e bromatológicas de amendoim forrageiro em diferentes intervalos de corte**. 2012. 96 p. Dissertação (Mestrado em Agronomia) – Universidade Federal do Acre, Rio Branco - Acre, 2012.

SANTOS, V. B. **Avaliação genotípica de linhagens de arroz de terras altas via metodologia de modelos mistos**. 2009. 152 f. Tese (Doutorado em Fitotecnia) – Universidade Federal de Lavras, Lavras, 2009.

SCARPINATI, E. A.; PERECIN, D.; CÉSAR DE PAULA, R. BONINE, C. A. V.; PAVAN, B. E.; CANDIDO, L. S. Influência do modelo de análise estatística e das formas das parcelas experimentais na seleção de clones de *Eucalyptus* spp. **Revista Árvore**, v. 33, n. 4, p. 769-776, 2009.

SCHENKEY, F. S. **Studies on effects of parental selection on estimation of genetic parameters on breeding values of metric traits**. University of Ghelph, Ghelph, Canada. (PhD. Thesis). 1998.

SEARLE, S. R.; CASELLA, G.; MCCULLOCH, C. E. **Variance components**. New York: J. Wiley, 1992. 528 p.

SEARLE, S.R. **Linear models**. New York: John Wiley, 1971. 532p.

SHIMOYA, A.; PEREIRA, A. V.; FERREIRA, R. de P.; CRUZ, C.;CARNEIRO, P. C. S. Repetibilidade de características forrageiras do capim-elefante. **Scientia Agricola**, v.59, n.2, p.227-234, 2002.

SILVA, H. D.; FERREIRA, D. F.; PACHECO, C. A. P. Avaliação de quatro alternativas de análise de experimentos em látice quadrado, quanto à estimação de componentes de variância. **Bragantia**, v. 59, n. 1, p. 117-123, 2000.

SMITH, A. B. C.; GILMOUR, A. The analysis of crop variety evaluation data in Australia. **Journal of Statistics**, Australia New Zealand, n. 43, p. 129-145. 2001.

SOUZA SOBRINHO, F.; AUAD, A. M.; LÉDO, F. J. da S. Genétic variability in *Brachiaria ruziziensis* for resistance to spittlebugs. **Crop Breeding and Applied Biotechnology**, n. 10, p. 83 – 88, 2010.

SOUZA SOBRINHO, F.; BORGES, V. LÉDO, F. J. da S.; KOPP, M. M. Repetibilidade de características agrônômicas e número de cortes necessários para a seleção de *Urochloa ruziziensis*. **Pesquisa Agropecuária Brasileira**, v. 45, n. 6, p. 579-584, 2010.

STATISTICAL ANALISYS SYSTEM INSTITUTE – SAS. **The SAS-System for Windows**. User's Guide. versão 9.1.3. Cary: Statistical Analysis System Institute, 2003.

STEEL, R. G. D.; TORRIE, J. H. **Principles and procedures of statistics**. 2 ed. Mc Graw-Hill, New York, 1980. 633 p

STEIN,C. Inadmissibility of the usual estimator for the mean of multivariate normal distribution. In Symposium on Mathematical Statistics and Probability, 3., 1955, Berkeley. **Proceedings...**Lincoln: University of Berkeley, 1955. P. 197-206.

STURION, J. A.; RESENDE, M. D. V. de R. Avaliação genética e análise de *deviance* em um teste desbalanceado de procedência e progênie de *Ilex paraguariensis*. **Pesquisa Florental Brasileira**, v. 30, n. 62, p. 157-160, 2010.

THOMAS, R.J. Role of legumes in providing N for sustainable tropical pasture legumes. **Plant and Soil**, v.174, n.1-2, p.103-118, 1995.

TOMÉ, M. A. **Modelagem do controle local para ensaios em látice aplicados ao melhoramento vegetal**. 2002 92 p. Dissertação (Mestrado) – Universidade Federal de Lavras, Lavras, 2002.

UNITED STATES DEPARTMENT OF AGRICULTURE – USDA. **World markets and trade**. Washington, USDA. Apr. 2013. Disponível em: <[http://www.fas.usda.gov/psdonline/circulars/livestock\\_poultry.pdf](http://www.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf)>. Acesso em: nov. 2013.

VALENTIM, J. F.; ANDRADE, C. M. S. de; MENDONÇA, H. A. de; SALES, M. F. L. Velocidade de estabelecimento de acessos de amendoim forrageiro na Amazônia Ocidental. **Revista Brasileira de Zootecnia**, v. 32, n. 6, p. 1569-1577, 2003.

VALENTIM, J. F.; CARNEIRO, J. da C.; SALES, M. F. L. **Amendoim forrageiro cv. Belmonte: leguminosa para a diversificação das pastagens e conservação do solo no Acre**. Embrapa Acre, 2001. 18 p. (Circular Técnica, 43).

VALLE, C. B. Genetic resources for tropical areas: Achievements and perspectives. In: 19<sup>o</sup> International Grassland Congress, 2001, Piracicaba. **Poceedings...** Lincoln: Fundação de Estudos Agrários Luiz Queiroz, 2001. P. 477- 482.

VALLE, C. B.; JANK, L.; RESENDE, R. M. S.; BONATO, A. L. V. Lançamento de cultivares forrageiras: o processo e seus resultados – cvs. Massai, Pojuca, Campo Grande, Xaraés. In: EVANGELISTA, A. R.; REIS, S. T.; GOMIDE, E. M. (Ed.). **Forragicultura e pastagens: temas em evidência – Sustentabilidade.** Lavras: UFLA, 2003. P. 179-225.

VALLS, J. F. M.; SIMPSON, C. E. Taxonomy, natural distribution, and attributes of *Arachis*. In: KERRIDGE, P. C., HARDY, B. (Ed.). **Biology and agronomy of forage Arachis**, Cali, Colômbia: Centro Internacional de Agricultura Tropical (CIAT), 1994. Chapter 1. p. 1-18.

VIANA, J. M. S.; FARIA, V. R.; SILVA, F. F. e; RESENDE, M. D. V. de. Combined selection of progeny in crop breeding using best linear unbiased prediction. **Canadian Journal of Plant Science**, n. 92, p. 553-562, 2012.

VILELA, L.; MARTHA Jr., G. B.; BARIONE, L. G. Adubação na recuperação e na intensificação da produção animal em pastagens. In: SIMPÓSIO SOBRE O MANEJO DA PASTAGEM, 21., Piracicaba, 2004. **Anais...** Piracicaba: FEALQ, 2004. p. 425-472.

XAVIER, L. H. **Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação.** 2000. 91 p. Dissertação (Mestrado em Agronomia) – Universidade de São Paulo, Escola Superior de Agricultura Luiz Queiroz, Piracicaba, 2000.

XAVIER, L. H.; DIAS, C. T. S. Acurácia do modelo univariado para análise de medidas repetidas por simulação multidimensional. **Scientia Agricola**, v.58, n.2, p.241-250, 2001.

ZENI-NETO, H.; OLIVEIRA, R. A. de; DAROS, E.; BESPALHOK-FILHO, J. 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, n. 4, p. 425-430, 2008.