



JALES MENDES OLIVEIRA FONSECA

**AVALIAÇÃO VISUAL EM BRAQUIÁRIA: NÚMERO DE
AVALIADORES E MODELOS DE ANÁLISE**

**LAVRAS - MG
2017**

JALES MENDES OLIVEIRA FONSECA

**AVALIAÇÃO VISUAL EM BRAQUIÁRIA: NÚMERO DE AVALIADORES E
MODELOS DE ANÁLISE**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

Prof. Dr. José Airton Rodrigues Nunes
Orientador

Dr. Fausto Souza Sobrinho
Coorientador

**LAVRAS- MG
2017**

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

Fonseca, Jales Mendes Oliveira.

Avaliação visual em braquiária: número de avaliadores e
modelos de análise / Jales Mendes Oliveira Fonseca. - 2017.
83 p. : il.

Orientador: José Airton Rodrigues Nunes.

Coorientador: Fausto Souza Sobrinho.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2017.

Bibliografia.

1. Avaliação visual. 2. Número de avaliadores. 3. Análise de
notas. I. Nunes, José Airton Rodrigues. II. Sobrinho, Fausto Souza.
III. Título.

JALES MENDES OLIVEIRA FONSECA

**AVALIAÇÃO VISUAL EM BRAQUIÁRIA: NÚMERO DE AVALIADORES E
MODELOS DE ANÁLISE**

**VISUAL EVALUATION IN BRACHIARIA: NUMBER OF EVALUATORS AND
MODELS ANALYSIS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

APROVADA em 23 de fevereiro de 2017.

Dr. João Luiz da Silva Filho

Embrapa/Arroz e Feijão

Dr. Júlio Silvio de Souza Bueno Filho

UFLA

Prof. Dr. José Airton Rodrigues Nunes
Orientador

Dr. Fausto Souza Sobrinho
Coorientador

**LAVRAS- MG
2017**

À memória de minha mãe Eloisa Fraga de Oliveira Fonseca,
DEDICO

AGRADECIMENTOS

A DEUS.

Ao meu pai Márcio Fonseca e à minha irmã Thayze Fonseca por estarem presentes em todos os momentos difíceis dessa jornada, dando-me apoio e incentivo diariamente.

Ao professor José Airton pela orientação precisa (e acurada), conhecimentos compartilhados, paciência, por ser um professor competente que sempre soube sanar dúvidas de âmbito acadêmico e profissional e por proporcionar palavras amigas nos momentos de maior estresse.

À minha namorada, meu amor, por ser a minha melhor companhia, ser carinhosa, estar sempre presente trazendo alegria aos meus dias, por tornar essa caminhada menos pesada e por conseguir me acalmar nos momentos de ansiedade, sou eternamente grato.

À família Paióça pelo convívio diário nos almoços comandados pela Michelinha, pelas festas que sempre rendiam boas risadas, pelo companheirismo e, principalmente, pela amizade de todos os moradores, muito obrigado.

Aos companheiros e companheiras do G-sorgo por tornarem os dias de campo em momentos de descontração.

Aos professores do Departamento de Biologia por serem profissionais exemplares.

Aos colegas do Gen por terem sido muito atenciosos desde o primeiro dia no programa de Genética e Melhoramento de Plantas.

À todos os profissionais do Departamento que sempre foram muito solícitos em todas as situações.

Ao professor Júlio Bueno pelo auxílio com as análises estatísticas bayesianas.

Ao Dr. Fausto e à Embrapa Gado de Leite pela confiança e disponibilidade do dados que deram origem a esse trabalho.

Ao CNPq pela bolsa de estudos.

À todos que contribuíram de forma direta ou indireta para a conclusão deste trabalho.

*“Demore o tempo que for para decidir o que você quer da vida,
e depois que decidir não recue ante nenhum pretexto,
porque o mundo tentará te dissuadir”.*

Friedrich Nietzsche

RESUMO

Os programas de melhoramento genético da *Urochloa ruziziensis* (R. Germ. & CM Evrard) Crins. (sin. *Brachiaria ruziziensis*) têm adotado a sensibilidade visual como parâmetro avaliativo dos genótipos. Devido ao desconhecimento do número adequado de avaliadores no campo e a maneira de proceder as análises das notas atribuídas aos genótipos, no presente trabalho objetivou-se determinar um número otimizado de avaliadores e apresentar diferentes alternativas para analisar notas. Para isso, foram avaliados 2204 genótipos de *U. ruziziensis* no delineamento em blocos aumentados utilizando duas testemunhas: Marandu (*U. brizantha*) e Basilisk (*U. decumbens*). Foi mensurada a produção da massa verde e atribuído notas de vigor associada à produtividade aos genótipos pela inspeção visual efetuada por seis avaliadores. Para a determinação do número de avaliadores, os dados foram analisados usando a abordagem de modelos mistos com recuperação de informação interblocos e intergenotípica considerando a nota de um avaliador e as médias das combinações de 2, 3, 4, 5 e 6 avaliadores. Para apresentar as diferentes alternativas de análise das notas foi empregado as abordagens de modelos lineares mistos na escala original das notas, modelos lineares mistos na escala da nota transformada via Box-Cox, modelos lineares mistos generalizados e modelos lineares mistos generalizados bayesianos. Para validação das análises, em ambos os casos, foram considerados parâmetros genéticos e experimentais relevantes aos programas de melhoramento auxiliados pelo emprego da avaliação visual. Com base nos resultados, foi possível concluir que a seleção indireta de genótipos de *U. ruziziensis*, por meio das notas de vigor, é válida e que o número otimizado de avaliadores para proceder as avaliações visuais é três. Além disso, não foi possível detectar diferenças significativas entre as alternativas de análise das notas, porém foi possível observar a adequação de se utilizar os modelos que atentem para a escala original da nota.

Palavras-chave: Avaliação visual, Número de avaliadores, Acurácia seletiva, Correlação, Herdabilidade, Melhoramento de forrageiras, Modelo Linear Misto Generalizado, Transformação Box-Cox, Inferência Bayesiana

ABSTRACT

Urochloa ruziziensis (R. Germ. & CM Evrard) Crins. (sin. *Brachiaria ruziziensis*) breeding programs have adopted visual sensitivity as an evaluation parameter of genotypes. Due to the lack of knowledge regarding the adequate number of evaluators in the field and the way of performing the analysis of grades attributed to genotypes, the objective of this study was to determine an optimized number of evaluators and present different alternatives for analyzing grades. In order to do that, 2204 *U. ruziziensis* genotypes were evaluated in augmented block design using two controls: Marandu (*U. brizantha*) and Basilisk (*U. decumbens*). Green mass production was measured and six evaluator attributed vigor grades based on visual inspection. To determine the number of evaluators, data were analyzed using mixed models approach with recovery of interblock and intergenotypic information considering grades given by one evaluator and by the average of 2, 3, 4, 5 and 6 evaluator combinations. In order to present the different alternatives for grade analysis, it was used approaches based on linear mixed models in original score scale, linear mixed models in Box-Cox transformed scale, generalized linear mixed models and Bayesian generalized linear mixed models. To validate the analysis, both cases considered genetic and experimental parameters used in breeding programs assisted by visual evaluation. Based on the results, it was possible to conclude that indirect selection of *U. ruziziensis* genotypes using grades is valid and that the optimal number of evaluators to carry out visual evaluations is three. Moreover, it was not possible to detect significant differences between the alternatives grade analysis, but it was possible to observe the adequacy offered by the usage of models that consider original grade scale.

Keywords: Visual assessment, Number of evaluators, Accuracy, Correlation, Heritability, Forage breeding, Generalized Linear Mixed Model, Box-Cox Transformation, Bayesian Inference

SUMÁRIO

1	INTRODUÇÃO GERAL	10
2	REFERENCIAL TEÓRICO	13
2.1	Melhoramento genético de <i>Urochloa ruziziensis</i>	13
2.2	Avaliação visual no melhoramento de plantas.....	16
2.3	Alternativas para análise de notas	20
2.3.1	Modelos lineares mistos.....	21
2.3.2	Transformação Box-Cox.....	25
2.3.3	Modelos lineares mistos generalizados	28
2.3.1	Modelos lineares mistos generalizados bayesianos.....	31
3	CONSIDERAÇÕES GERAIS	35
	ARTIGO 1 – NÚMERO DE AVALIADORES NA INSPEÇÃO VISUAL DA PRODUÇÃO DE MASSA VERDE EM <i>Urochloa ruziziensis</i>.....	44
	ARTIGO 2 – MODELOS DE ANÁLISE DE NOTAS NO MELHORAMENTO DE <i>Urochloa ruziziensis</i>.....	55
	APÊNDICE A – Rotina	78

PRIMEIRA PARTE

1. INTRODUÇÃO GERAL

A *Urochloa ruziziensis* (R. Germ. & CM Evrard) Crins. (sin. *Brachiaria ruziziensis*) vem despertando interesse entre os pesquisadores envolvidos no processo de melhoramento genético de culturas forrageiras devido ao seu potencial agronômico e, sobretudo, nutricional (VALLE et al. 2000; LOPES et al. 2010; SOUZA SOBRINHO et al. 2010a), sua capacidade de participar de cruzamentos intra e interespecíficos (NDIKUMANA, 1985; SAVIDAN; JANK; VALLE, 1989; TIMBÓ et al. 2014) e sua importância quando se deseja avançar gerações em programas de melhoramento que optem por realizar o método de seleção recorrente (VALLE et al. 2009). Por essas razões, melhoristas têm buscado métodos para avaliar genótipos de *U. ruziziensis* de forma eficaz, possibilitando a seleção dos genótipos superiores.

Dentre os métodos adotados, aqueles que se utilizam de escalas diagramáticas (notas) ganham destaque quando se deseja avaliar genótipos em etapas iniciais de ciclos de seleção em um programa de melhoramento genético. Isso ocorre porque nessas etapas há um grande número de genótipos a serem avaliados, o que dificulta uma fenotipagem adequada. No que concerne ao melhoramento de espécies forrageiras, características como resistência à pragas e doenças (SILVA, 2013), densidade de perfilho (FIGUEIREDO; NUNES; VALLE, 2012), florescimento, hábito de crescimento (NGUYEN; SLEPER, 1983), rebrota, tolerância ao frio (FÈ, 2015), “stay green” (WILKINSON, 2003), vigor da planta (TEIXEIRA, 2015) e produção de massa verde (PMV) (BURTON, 1974; BURTON, 1982; PRICE; CASLER, 2014) podem ser avaliadas por meio de notas. Contudo, para que essas avaliações visuais sejam eficazes, preconiza-se que as características avaliadas possuam alta correlação positiva com a sensibilidade visual do avaliador e que o caráter em questão possua herdabilidade relativamente alta, permitindo afirmar que a maior parte da expressão fenotípica se dá devido a variações genéticas, isto é, grande parte do potencial observado da planta seja herdável (BERNARDO, 2010).

O emprego da avaliação visual gera alguns questionamentos dos quais dois se destacam devido a frequência com que ocorrem. O primeiro está relacionado ao número de avaliadores visuais necessário para obtenção de dados suficientes que concedam, simultaneamente, fidedignidade às estimativas paramétricas e viabilidade econômica ao experimento. Enquanto a adoção de um número pequeno de avaliadores possa ser criticado

devido a subjetividade inerente à inspeção visual, podendo impactar negativamente nas estimativas dos parâmetros, a adoção de um número grande também está passível de críticas, devido ao elevado aporte de recursos demandados para contratá-los. Assim, fica evidente a necessidade de realizar um estudo que estime um número otimizado de avaliadores, de modo que a adoção de uma determinada quantidade passe a ser amparada por fundamentos técnicos. O segundo questionamento é concernente a maneira de se analisar a nota, visto que sua escala é diferente daquelas classificadas como variáveis contínuas.

A escala de notas é classificada como uma variável discreta ordinal (AGRESTI, 2002), conseqüentemente, sua análise por meio de modelos que pressupõem distribuição normal, por apoiar-se no teorema do limite central (MOOD; GRAYBILL; BOES, 1974), ou modelos que adotem uma mudança na escala da variável (BOX; COX, 1964), pode remeter à violação do pressuposto de normalidade. Dessa forma, o pesquisador fica sujeito a cometer erros ao tecer conclusões fundamentadas nos resultados experimentais. Mesmo assim, abordagens estatísticas baseadas na distribuição normal ou em sua aproximação têm sido frequentemente utilizadas pelos melhoristas.

A adoção de modelos inadequados para análise dos dados pode implicar na estimação de parâmetros que não representam a verdade, produzindo resultados espúrios (JAEGER, 2008; BOLKER, 2009; STROUP, 2015). No contexto do melhoramento genético de plantas, tal inadequação pode culminar em um ordenamento errado dos genótipos superiores (erro tipo 3), permitindo a eliminação de genótipos promissores, bem como a seleção de genótipos de menor performance.

Alguns modelos para analisar dados que possuam distribuição diferente da gaussiana, a exemplo dos modelos lineares mistos generalizados (MLMG), surgiram com o refinamento de técnicas estatísticas e desenvolvimento de programas computacionais robustos. Estes modelos são uma extensão do modelo linear misto (MLM) por englobar distribuições da família exponencial, da qual a multinomial é frequentemente utilizada para descrever variáveis categóricas ordinais (e.g. notas). Modelos multinomiais podem ser empregados sob enfoque frequentista (BRESLOW, CLAYTON, 1993; MCCULLOCH, 1997) ou bayesiano (GIANOLA & FOULLEY, 1983; GIANOLA & FERNANDO, 1986). Quando se utiliza os modelos lineares mistos generalizados bayesianos (MLMGB) as estimativas paramétricas são revisadas pelo teorema de Bayes (BAYES; PRICE, 1763), o qual afirma que a probabilidade *a posteriori* da ocorrência de um evento é proporcional a verossimilhança da ocorrência desse evento ponderado por um conhecimento *a priori*.

Diante do exposto, no presente trabalho objetivou-se determinar um número otimizado de avaliadores no campo para o procedimento da avaliação visual no melhoramento genético da *U. ruziziensis*. Analisar dados provenientes da inspeção visual utilizando quatro alternativas de análise: abordagem de modelos lineares mistos na escala original das notas, modelo linear misto na escala da nota transformada via Box-Cox, modelo linear misto generalizado e modelo linear misto generalizado bayesiano. Objetivou-se também analisar os dados provenientes da PMV dos genótipos utilizando um modelo linear misto com a finalidade de proceder ao estudo da correspondência dos genótipos selecionados por meio das notas, para verificar se a avaliação por notas possibilita selecionar os genótipos que sejam realmente superiores.

Devido à produção de dois trabalhos distintos, optou-se por apresentar a presente dissertação no formato de artigo. Assim, a primeira parte foi composta por uma introdução geral, um referencial teórico, reportando assuntos de ambos os trabalhos e um relato das contribuições deste trabalho. A segunda parte foi constituída, propriamente, pelos dois trabalhos no formato de artigo, respeitando as normas das revistas as quais serão submetidos. O primeiro artigo contemplou a determinação do número otimizado de avaliadores no campo e o segundo apresentou alternativas de análise das notas.

2. REFERENCIAL TEÓRICO

2.1. Melhoramento genético de *Urochloa ruziziensis*

A *Urochloa ruziziensis* (R. Germ. & CM Evrard) Crins. (Sinonímia. *Brachiaria ruziziensis*) é uma espécie africana nativa do Valle Ruzi no Zaire e Burundi (RENVOIZE; CLAYTON; KABUYE, 1996). Foi amplamente distribuída nos países tropicais onde ficou conhecida como ruziziensis ou capim Congo e, em inglês, “Congo signalgrass” e “Kennedy ruzigrass” (MAASS, 1996). Possui aproximadamente 1,0 m de altura, com rizomas curtos, robustos e globosos. Colmos geniculados, decumbentes e radicantes nos nós inferiores. Possui inflorescência em panícula com racemos bilaterais, terminais, de 15 a 25 cm de comprimento com três a sete ramificações primárias, ascendentes, arqueadas, alternas distantes ente si, terminado em uma espiguetta rudimentar. Os ramos são de 4 a 10 cm de comprimento com raque alada, de 5,0 mm de largura, envolvendo a base da espiguetta. Seu florescimento é sincronizado e concentrado no final do verão (VALLE et al., 2013). Pode ser propagada tanto por sementes como vegetativamente.

O melhoramento genético do gênero *Urochloa* P. Beauv. (sin. *Brachiaria* (Trin.) Griseb.) é uma atividade recente, iniciada no Brasil em meados da década 1980 com a introdução de germoplasma exótico (VALLE, 2009). Desde então, a seleção de genótipos superiores, a partir da variabilidade genética natural existente, vem sendo o principal método de desenvolvimento de cultivares utilizado pelas empresas públicas nacionais, principais instituições responsáveis por promover o melhoramento genético de forrageiras no país (MILES; VALLE, 1996; HACKER; JANK, 1998; PEREIRA et al., 2001; JANK et al., 2005; MILES, 2006; SOUZA SOBRINHO, 2010a). Contudo, este método é limitado por basear-se somente na avaliação da capacidade adaptativa de acessos coletados na natureza, isto é, não gera novas combinações genéticas que ampliam a variabilidade da população, caracterizando-o como um método finito.

A *Urochloa ruziziensis* teve seu destaque nesse contexto por ser a única espécie do gênero exclusivamente sexual e diploide ($2x = 2n = 18$) (FERGUSON; CROWDER, 1974), permitindo de forma facilitada, a ampliação da variabilidade genética por meio da recombinação intra e interespecífica e posterior seleção de genótipos superiores (SOUZA SOBRINHO et al., 2005). A recombinação interespecífica só foi possível após a duplicação do seu número cromossômico utilizando colchicina (GOBBE et al. 1981; SWENNE et al. 1981), de modo que a ploidia fosse igualada à das espécies mais importantes do gênero (*U.*

decumbens e *U. brizantha*) viabilizando a hibridação entre elas (NDIKUMANA, 1985; SAVIDAN; JANK; VALLE, 1989). As cultivares Mulato I e Mulato II, lançadas pelo CIAT, são exemplos de sucesso da aplicação dessa técnica (SOUZA SOBRINHO et al., 2009). Timbó et al. (2014), utilizando a mesma técnica para duplicação de cromossomos, concluíram que genótipos de *U. ruziziensis*, já adaptados às condições ambientais brasileiras, produzem pólen férteis em porcentagens que tornam viáveis os cruzamentos em programas de melhoramento genético. Pereira et al. (2005) avaliaram 47 híbridos interespecíficos originados do cruzamento entre *U. ruziziensis* x *U. brizantha* e *U. ruziziensis* x *U. decumbens* com o objetivo de selecionar novas cultivares melhoradas que combinassem alto rendimento e qualidade de forragem, de modo a intensificar a produção de leite sob pastejo. Dados como produção total de matéria seca, porcentagem de matéria seca, altura de planta e vigor de rebrota de seis cortes foram coletados. Amostras do terceiro corte foram analisadas por valor nutritivo usando espectroscopia no infravermelho próximo (NIRS). Os resultados mostraram a existência de variabilidade para a maioria das características analisadas, confirmando que é possível realizar a seleção de híbridos interespecíficos que combine alta qualidade e alta produtividade.

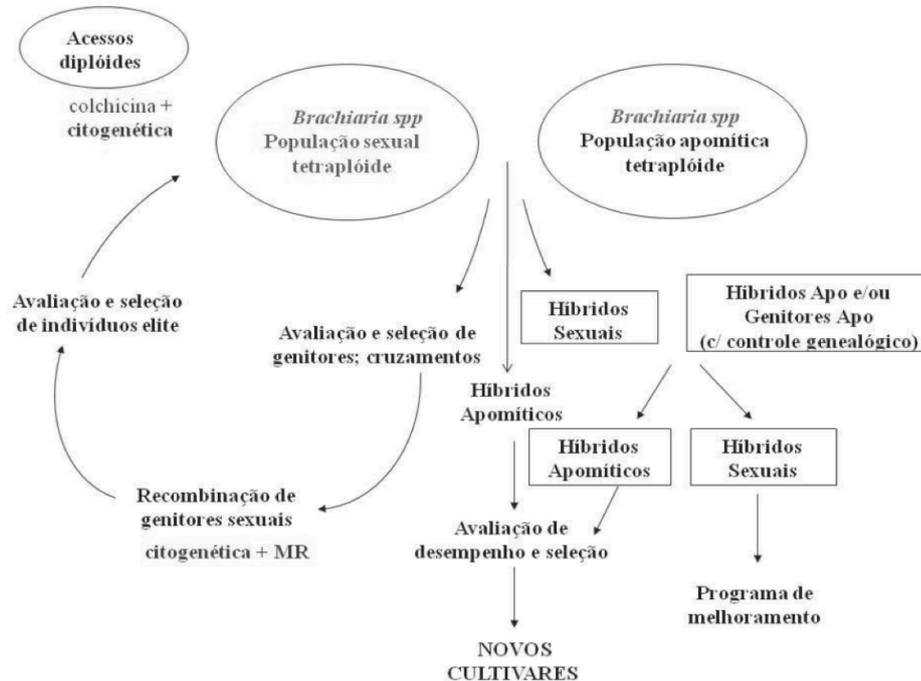
Além de sua capacidade de participar de cruzamentos intra e interespecíficos, a *U. ruziziensis* se destaca também pelas suas características agrônomicas e, sobretudo, nutricionais. Ela se estabelece rapidamente e apresenta grande crescimento no início da estação chuvosa, possui compatibilidade com leguminosas e seu florescimento é concentrado, o que favorece a produção de uma grande quantidade de sementes. A espécie também é uma boa alternativa para o plantio direto pela rapidez de estabelecimento, é dotada de elevada qualidade de massa, tem facilidade de dessecação por herbicidas (VALLE et al. 2000) e boa tolerância ao sombreamento, conferindo a ela aptidão para o emprego em sistemas de integração lavoura-pecuária e floresta (ILPF). Lopes et al. (2010) conduziram um experimento com o objetivo de avaliar a digestibilidade ruminal *in situ* da matéria seca, proteína bruta e fibra em detergente neutro da forragem de quatro espécies de *Urochloa* (*U. brizantha*, *U. humidicola*, *U. decumbens*, *U. ruziziensis*). Observaram que a *U. ruziziensis* apresentou o maior valor nutricional, as mais elevadas taxas de degradação e degradabilidades efetivas da matéria seca e da proteína bruta e o menor teor de fibra em detergente neutro associado à maior taxa de degradação desse nutriente. Com relação à aceitação pelo rebanho bovino, o autor menciona que a *U. ruziziensis* apresenta preferência quando comparada a outras braquiárias, devido a sua qualidade acentuada e palatabilidade distinta. Serrão e Simão

Neto (1971) consideram a aceitação de *U. decumbens* e *U. ruzizensis* como muito boa para o gado, ainda que as plantas apresentem estágio avançado de maturação.

Em contrapartida, a *U. ruzizensis* apresenta algumas características agronômicas que dificulta seu cultivo nas diversas regiões de pasto brasileiro. Por isso, algumas destas características têm sido alvo do melhoramento genético da espécie. A *U. ruzizensis* apresenta produtividade mais baixa, é pouco tolerante a solos ácidos e exige grande quantidade de nutrientes, logo é recomendada para solos bem drenados de média a alta fertilidade com clima de região tropical. A espécie não resiste à geada e apresenta alta susceptibilidade às cigarrinha-das-pastagens. Contudo, Miguel et al. (2011) ao avaliarem a variabilidade genética da *U. ruzizensis* para a tolerância ao alumínio, concluíram que é possível selecionar genótipos tolerantes a solos ácidos, visto que existe variabilidade na espécie para este caráter. Souza Sobrinho et al. (2010b) avaliaram a resistência de *U. ruzizensis* à cigarrinha-das-pastagens, espécies *Deois schach* e *Mahanarva spectabilis*. Com isso, observaram variabilidade para o caráter, permitindo a identificação de genótipos resistentes.

A *U. ruzizensis* também é utilizada como ferramenta auxiliar na condução de um programa de melhoramento genético de forrageiras. A Embrapa Gado de Corte utiliza a *U. ruzizensis* tetraploidizada estrategicamente para avançar os ciclos de seleção recorrente. Em um primeiro momento, a *U. ruzizensis* é inter cruzada gerando genótipos elite que são utilizados em cruzamentos interespecíficos com espécies apomíticas facultativas. Em seguida, os híbridos apomíticos são avaliados em ensaios VCU a fim de que uma nova cultivar seja lançada. Concomitantemente, os híbridos sexuais que vão surgindo são também avaliados e aqueles que apresentarem melhor performance são adicionados ao programa de melhoramento genético para participarem de novos cruzamentos (VALLE et al. 2009). A figura 1 apresenta o esquema de melhoramento genético para a obtenção de cultivares de *Urochloa spp.* por hibridação interpopulacional utilizada na Embrapa gado de corte. É importante ressaltar que a adoção da seleção recorrente entre espécies diferentes de *Urochloa* possibilita o acúmulo de alelos de resistência à determinadas pragas. Miles, Cardona e Sotelo (2006) realizaram um estudo em populações sintéticas do gênero *Urochloa* composta por três espécies diferentes: *U. ruzizensis*, *U. decumbens* e *U. brizantha*. Os autores concluíram que os cruzamentos interpopulacionais propiciaram a formação de genótipos com antibiose a três espécies de cigarrinha (*Aeneolamia varia*, *A. reducta* e *Zulia carbonaria*), já que os genótipos inviabilizaram o desenvolvimento de ninfas. O trabalho reforça também a possibilidade de obtenção de clones sexuais tetraplóides de *U. ruzizensis* com níveis elevados de resistência às espécies citadas.

Figura 1. Esquema de melhoramento para a obtenção de cultivares de *Urochloa spp.* (= *Brachiaria spp.*) por hibridação interpopulacional.



Fonte: Valle et al. (2009).

A Embrapa gado de leite é responsável pelo programa brasileiro de melhoramento genético de *U. ruziziensis* (JANK et al., 2014). O programa tem como objetivo a obtenção de cultivares com maior resistência às cigarrinhas, maior tolerância a solos de baixa fertilidade e alta produção de biomassa verde e seca (SOUZA SOBRINHO et al., 2009). Para isso, cabe aos melhoristas adotarem estratégias economicamente viáveis e, sobretudo, eficazes que assegurem o cumprimento de tais objetivos. Uma dessas estratégias é a utilização da avaliação visual nas diversas etapas de ciclo de seleção, a qual tem se demonstrado satisfatória.

2.2. Avaliação visual no melhoramento de plantas

A avaliação visual é a técnica mais antiga utilizada pelo homem na seleção de plantas. Sua origem data de mais de 10 mil anos, coincidindo com o surgimento da agricultura e consequente domesticação das plantas (HALLAUER, 2011). Ela permite que a seleção fenotípica seja realizada de forma criteriosa, sem que haja escolha de plantas por mero acaso (TEIXEIRA, 2015). No início, a seleção fenotípica (ou visual) visava identificar as plantas produtoras de alimentos que proveria recursos às civilizações primitivas. Todavia, desde os

anos 1900, essa seleção fenotípica, intitulada comumente de seleção massal ou massal estratificada, passou a ser adotada como uma estratégia de melhoramento genético, utilizada por diversos melhoristas e pesquisadores, em várias etapas de um programa de melhoramento visando o aperfeiçoamento de determinados caracteres alvo (SMITH, 1909; ATKINS, 1964; STUTHMAN; STEIDL, 1976; BURTON, 1974, 1982; BRUCKNER et al., 1991; ABREU, 2010; SILVA et al., 2013; PRICE; CASTLER, 2014; ABADASSI, 2014).

A avaliação visual consiste na atribuição de notas (ou escores) aos genótipos em teste, assim viabiliza, de forma prática, a discriminação destes, permitindo que o pesquisador selecione aqueles que apresentarem superioridade com relação às características em estudo. No que concerne ao melhoramento de espécies forrageiras, características como resistência à pragas e doenças (SILVA, 2013), densidade de perfilho (FIGUEIREDO; NUNES; VALLE, 2012), florescimento, hábito de crescimento (NGUYEN; SLEPER, 1983), rebrota, tolerância ao frio (FÈ, 2015), “stay green” (WILKINSON, 2003), produção de massa verde (PMV) (BURTON, 1974; BURTON, 1982; PRICE; CASLER, 2014) e vigor da planta são avaliadas por meio da atribuição de notas.

Um indicador da eficiência da avaliação visual é a herdabilidade (h^2). Ela indica a proporção da variância genética presente na variância fenotípica total. Portanto, a seleção visual terá maior eficácia quando os caracteres mensurados possuírem valores elevados de herdabilidade (BURTON, 1974, 1982; CUTRIM, 1997). Segundo Bernardo (2010), um baixo valor de herdabilidade indica que o fenótipo é um pobre indicador do valor reprodutivo, logo a seleção visual fica comprometida. Outro aspecto importante da herdabilidade é a sua relação com a acurácia seletiva (\hat{r}_{gg}) (RESENDE, 2007). Em condições experimentais balanceadas, a acurácia corresponde a raiz quadrada da herdabilidade (RESENDE, 2002). Porém, em condições desbalanceadas a acurácia corresponde à correlação entre o valor genético verdadeiro e os valores genéticos preditos a partir dos dados experimentos e é definida por meio da equação (HENDERSON, 1984): $\hat{r}_{gg} = (1 - PEV / \sigma_g^2)^{1/2}$, em que PEV corresponde a variância do erro de predição e σ_g^2 a variância genética entre os genótipos. Vale ressaltar que a acurácia é um parâmetro que remete a precisão experimental, logo quanto menor for a PEV, maior é a acurácia e, conseqüentemente, maior a precisão (RESENDE, 2007).

Os caracteres mais adequados para utilização da seleção visual são de caráter qualitativo, isto é, aqueles controlados por poucos genes e que sofrem menor influência do ambiente (ATKINS, 1964). Em contrapartida, os caracteres quantitativos, aqueles controlados por muitos genes e que sofrem maior influência do ambiente também podem ser estudados.

Em ambos os casos, é necessário que as técnicas de experimentação sejam primorosamente executadas, de modo que haja um controle eficaz do ambiente que permita ao pesquisador encontrar diferenças reais entre os genótipos.

Cutrim et al. (1997) avaliaram diversos aspectos relacionados a capacidade da seleção visual para identificar diferenças na produtividade de grãos de arroz irrigado entre indivíduos da geração F₂. Dentre os resultados obtidos, destacaram a baixa eficiência da seleção visual nesse processo. Resultados semelhantes foram obtidos por Abreu et al. (2010) e Abadassi (2014) ao conduzirem um experimento com a cultura do milho. Os autores concluíram que, mesmo com a estratificação da área, a seleção visual continuava ineficiente para selecionar genótipos de melhor performance para produtividade de grãos. Contudo, Bruckner et al. (1991), ao analisarem plantas de centeio aptas a produção de forragem, constataram que a seleção visual foi efetiva para identificar plantas de elevada produtividade. Os pesquisadores ressaltaram que o mérito da seleção se deu devido a alta correlação positiva entre a classificação visual das plantas e seus pesos, tendo registrado correlações de 0,82 até 0,90. Magnitudes similares foram apresentadas por Riday (2009) ao estimar a correlação entre as notas associadas à biomassa e a produção de forragem de “red clover” (*Trifolium pratense* L.) em viveiros. O autor concluiu que a utilização das notas para seleção de plantas produtivas foi tão precisa quanto as aferições reais dos rendimentos.

Burton (1974) e Burton (1982) propuseram um novo método de seleção visual por meio da adoção de algumas restrições às avaliações visuais tradicionais (seleção massal) utilizando a espécie forrageira *Paspalum notatum* var. *saure* “Parodi”. No primeiro momento, o método ficou conhecido como “Recurrent Restricted Phenotypic Selection” e preconizava quatro restrições: selecionar os cinco melhores genótipos de cada talhão formado por 5x5 plantas, manter 200 genótipos da população para o ciclo seguinte, inter cruzar os fenótipos selecionados para o próximo ciclo, isoladamente, em laboratório e inter cruzar três inflorescências por fenótipo selecionado. Posteriormente, o autor incrementou outras condições ao método e o intitulou de “Improved Recurrent Restricted Phenotypic Selection”. Dentre estas, ressaltou a importância de se registrar o parentesco das progênes e o cuidado necessário na execução dos tratos culturais, enfatizando a importância da uniformidade experimental para a seleção visual. Segundo o autor, o método proporciona ganhos elevados ao longo dos ciclos, além de reduzir custo e tempo na execução das atividades.

A literatura, portanto, nos leva a crer que a seleção visual pode ser considerada um método eficaz quando as características avaliadas possuírem alta correlação positiva com a sensibilidade visual do avaliador. Além disso, é importante ressaltar que, embora a correlação

e/ou herdabilidade não apresente magnitudes adequadas que justifique a utilização da seleção visual para selecionar genótipos de alta performance, alguns pesquisadores destacam sua prática para realização do descarte de genótipos indesejáveis (ATKINS, 1964; BURTON, 1982).

O programa de melhoramento genético da *U. ruziziensis*, conduzido pela Embrapa Gado de Leite, tem adotado a avaliação visual nas etapas iniciais dos ciclos de seleção do programa de melhoramento genético. A opção por essa técnica é justificável devido a grande quantidade de genótipos existentes a serem avaliados nessas etapas, o elevado aporte de recursos demandados para realizar uma fenotipagem adequada das plantas, a exemplo do custo com mão-de-obra para corte e pesagem de cada parcela, e o tempo gasto para que toda essas operações sejam realizadas. O método consiste na atribuição de notas considerando o vigor associado ao aspecto produtivo da planta, sendo adotado uma escala de 1 (baixa produção) a 5 (alta produção). Silva et al. (2013) adotaram esse método para verificar a resistência de clones da *U. ruziziensis* à *Collaria oleosa*. Para isso, conduziram um experimento em casa de vegetação utilizando 80 clones já infestados pela praga previamente. As notas advindas da porcentagem das injúrias causadas pela *C. oleosa*, bem como a redução dos teores de clorofila foram avaliados. Os autores observaram a existência de correlação positiva para as variáveis avaliadas, mostrando que é possível selecionar plantas resistentes de forma indireta. Ademais, encontraram valores acima dos 50% para herdabilidade média em ambas as características, o que evidencia eficiência na seleção de *U. ruziziensis* para o caráter resistência a *C. oleosa*, além de indicar que tal resistência pode ser transferida às gerações futuras.

A eficiência da seleção visual em *U. ruziziensis* foi estimada por Teixeira (2015) ao comparar a seleção visual com outros três métodos de seleção: seleção massal baseada na média fenotípica, na média BLUP (Melhor Predição Linear Não-tendenciosa) e na escolha ao acaso de plantas, isto é, de forma aleatória. O parâmetro utilizado foi o número de coincidências entre as plantas com maiores médias de notas de vigor associada à produtividade e as demais estratégias de seleção. Para a comparação entre seleção visual e média fenotípica, observaram variações de 60% a 70% com intensidade de seleção variando de 4% a 14% respectivamente. Resultados inferiores foram observados na comparação entre seleção visual e seleção na média BLUP de 49% a 54% com a mesma variação na intensidade de seleção. Dessa forma, concluíram que a seleção visual pode ser considerada uma boa ferramenta em etapas iniciais dos ciclos de seleção em programas de melhoramento genético quando a intensidade de seleção não for muito acentuada.

Dois fatos presentes em todos os trabalhos citados nessa revisão merecem uma atenção especial. O primeiro está relacionado a variação do número de avaliadores responsáveis pela inspeção visual dos genótipos em campo. Enquanto alguns trabalhos adotaram somente um avaliador para atribuição das notas, outros adotaram dez (ABREU, et al., 2010). Portanto, é necessário determinar um número otimizado de avaliadores, com base em fundamentos técnicos, para que a adoção de uma certa quantidade de avaliadores não se dê arbitrariamente. O segundo está relacionado à maneira de analisar as notas atribuídas aos genótipos. Em todos os trabalhos, os autores assumiram que as notas seguiam distribuição normal, ainda que essas notas, no contexto de melhoramento genético, remetam ao conceito de delimitação de categoria. Logo, é pertinente pensar que o caráter nota faça parte da classe das variáveis categóricas e, por possuir níveis ordenados, é classificada como multinomial ordinal (AGRESTI, 2002). Dessa forma, assumir que as notas sigam distribuição normal pode conduzir a obtenção de resultados espúrios (JAEGER, 2008, BOLKER, 2009; STROUP, 2015). Portanto, é esperado que haja maior confiabilidade nas análises que adotem métodos condizentes com a realidade da escala das notas.

2.3. Alternativas para análise de notas

Um modelo estatístico fornece uma estrutura teórica para uma melhor compreensão de um fenômeno de interesse. Logo, pode ser definido como uma construção matemática que possibilita a representação do mecanismo que gera as observações. Embora o modelo postulado seja uma simplificação idealizada da complexidade do mundo real, ele pode fornecer aproximações úteis das relações entre as variáveis (RENCHE; SCHAALJE, 2008). Citando Box e Draper (1987), “all models are wrong, but some are useful” (“todos os modelos estão errados, mas alguns são úteis”). Portanto, é importante destacar que não existe um modelo perfeito que descreva as causas ou associações entre as variáveis adotadas para o estudo de um determinado fenômeno, mas existe um modelo adequado que permita obter conclusões razoáveis a respeito deste. Cabe ao profissional da área de melhoramento genético saber identificar um modelo que melhor se ajuste às características que ele julgar serem importantes na condução de seu programa. Assim, esse modelo subsidiará a compreensão dos fatores genéticos e ambientais que influenciam a expressão desses caracteres.

Esta revisão não pretendeu detalhar os desenvolvimentos estatísticos envolvidos em cada método abordado, porém apresentou, de maneira simplificada, alternativas de análise para dados de notas. Para maiores detalhes a respeito dos métodos é indicado: Searle (1971),

Rencher e Schaalje (2008), Searle, Casella e McCulloch (1992), Agresti (2002), Agresti (2007), McCullagh e Nelder (1989), McCulloch, Searle e Neuhaus (2008), Stroup (2012), Sorensen e Gianola (2002), Bosltad (2007), dentre outros.

2.3.1. Modelos lineares mistos

Um modelo linear misto, chamado sumariamente de modelo misto, é aquele em que coexistem fatores de efeitos fixos e aleatórios, além do erro experimental. A decisão sobre como tratar o efeito de um fator é amplamente referenciada na literatura (EISENHART, 1947; SEARLE, 1971; SEARLE, CASELLA E MCCULLOCH 1992; MCCULLOCH, SEARLE E NEUHAUS, 2008). Basicamente, o efeito de um fator é dito fixo quando o interesse do estudo se resume aos níveis adotados pelo pesquisador, logo os resultados advindos desse estudo não são extrapolados para outros níveis. Um exemplo claro de um fator tratado como fixo é a dosagem de um adubo numa área experimental onde os níveis são uma quantidade preestabelecida de Kg/ha do adubo. Assim, os efeitos da dosagem serão restritos aos níveis preestabelecidos. Em contrapartida, o efeito de um fator é tratado como aleatório quando os resultados de um estudo podem ser extrapolados para outros níveis. Portanto, é razoável pensar que o efeito aleatório nada mais é que uma realização eventual de um determinado fenômeno.

No contexto do melhoramento vegetal, o efeito de genótipo pode ser tratado como fixo ou aleatório dependendo do estágio em que o programa se encontra. Em estágios avançados, a exemplo dos ensaios de VCU, é coerente tratar o efeito de genótipo como fixo, pois o interesse do melhorista está inteiramente voltado para os genótipos em questão, uma vez que, nesse momento, o importante é a recomendação comercial daqueles melhores. Contudo, em estágios iniciais, o efeito de genótipo pode ser tratado como aleatório, já que é sensato pensar que essas progênies são um pequena amostra das possíveis progênies oriundas de um determinado cruzamento (NUNES, 2006). O presente trabalho trata o efeito de genótipo como aleatório, justificando, portanto, a opção pelos modelos mistos para o processamento de suas análises.

A abordagem de modelos mistos pode ser utilizada para especificar os consagrados modelos de regressão, análise de variância (ANAVA) e de análise de covariância (ANCOVA) (RENCHEER; SCHAALJE, 2008). No caso da ANAVA, método frequentemente empregado por melhoristas para execução de testes estatísticos (média e significância), além da

preocupação com a natureza dos efeitos dos fatores, faz-se necessário a validação de seus pressupostos, de modo que os resultados obtidos sejam profícuos (EISENHART, 1947).

A utilização dos métodos ANAVA para análise de notas proporcionou avanços significativos para o melhoramento de forrageiras, principalmente em estudos relacionados à resistência à pragas (CARDONA, 1999; SMITH, 2001; CARDONA, 2004; PABÓN, 2007; LÓPEZ, 2009; SOUZA SOBRINHO, 2009; CARDONA, 2010; AGUIRRE, 2013), justificando sua preferência pelos profissionais da área. Outro aspecto que alavancou a simpatia por esses métodos foi a praticidade das interpretações dos resultados e, sobretudo, as implementações matemáticas em softwares que facilitava o procedimento das análises. Contudo, toda essa praticidade perde sua prerrogativa em experimentos desbalanceados (SEARLE et al., 1992; MCCULLOCH et al., 2008), favorecendo o emprego de modelos menos restritivos que estimem parâmetros de forma robusta enquanto mantém um alto poder dos testes.

As limitações dos métodos ANAVA (JAEGER, 2008, BOLKER, 2009; STROUP, 2012, STROUP, 2015), juntamente com as alterações conjunturais da atualidade, a exemplo dos avanços na área da estatística, computacional e, principalmente, aos avanços tecnológicos, propiciaram a adoção de modelos mais flexíveis. Nesse contexto, o emprego de modelos mistos gerais vêm ganhando espaço na forma de proceder essas análises.

Um modelo misto geral é apresentado a seguir (HENDERSON, 1984):

$$y = X\beta + Za + e \quad (2.1)$$

em que:

y: vetor de notas, de dimensões $n \times 1$;

X: matriz do modelo referente aos efeitos fixos β , de dimensões $n \times p$;

β : vetor dos efeitos fixos, de dimensões $p \times 1$;

Z: matriz do modelo referente aos efeitos aleatórios a , de dimensões $n \times q$;

a: vetor dos efeitos aleatórios dos genótipos, com $a \sim NMV(0, G)$,

e: vetor de erros, de dimensões $n \times 1$, com $e \sim NMV(0, R)$.

O melhoramento genético, na perspectiva de modelos mistos, assume G como sendo a matriz de variância-covariância genética e pode incorporar um eventual parentesco (matriz A) existente entre as progênes envolvidas (HENDERSON, 1976; RESENDE, 2000; NUNES,

2008). Em contrapartida, R refere-se a matriz de variância-covariância dos resíduos e também pode ser modelada de modo a acomodar situações como a heterocedasticidade, evidenciando, dessa forma, a maior flexibilidade do modelo linear misto. Entretanto, G e R não são conhecidas, sendo as componentes de variância estimadas usando métodos como REML (residual/restrict maximum likelihood) (PATTERSON, THOMPSON, 1971; SEARLE CASELLA E MCCULLOCH, 1992). Vale ressaltar que, embora o modelo supracitado tenha sido definido para notas, o mesmo será utilizado para análise da PMV, logo, nesses caso, o vetor y será definido pelos pesos dos genótipos.

A abordagem de modelos mistos surgiu no melhoramento animal (HENDERSON, 1959; HENDERSON, 1984) e tem sido largamente empregada no melhoramento vegetal (RESENDE, 2000; PIEPHO et al., 2008; CROSSA, 2010). Segundo a literatura, sua aplicação para analisar dados de notas em espécies forrageiras tem-se intensificado nos últimos anos. Fè et al. (2015) empregaram o modelo misto para estimar a variância genética e ambiental presente num programa de melhoramento genético comercial de azevém (*Lolium perenne* L.) estabelecido em sete locais diferentes da Europa. Para isso, avaliaram, visualmente, seis características da referida espécie, sendo cinco delas por meio de uma escala de notas, a citar, resistência à ferrugem, rebrota, taxa de crescimento no período da primavera, densidade de planta e resistência ao inverno. Dentre os resultados, os autores destacam a existência de variabilidade genética significativa, além de elevada herdabilidade para todas as características, sendo a resistência à ferrugem e a rebrota aquelas de maior magnitude. Ressaltaram a presença de correlação entre as características e discutiram os motivos da significância para a interação genótipos por ambientes, indicando que, parte desta seja resultado da avaliação realizada com base na atribuição de notas.

Silva et al. (2013) avaliaram a variabilidade genética em clones de *U. ruziziensis* para a característica resistência à *Collaria oleosa* (Hemiptera: Miridae) baseando-se na intensidade de injúria foliar. Para isso, mantiveram 80 clones em ambiente protegido, previamente infestado pela praga, por 40 dias. Posteriormente, a redução do teor de clorofila e a porcentagem de injúria sofrido foram avaliados. Na avaliação da característica injúria da folha, utilizaram uma escala de notas respeitando os critérios apresentados por Cardona et al. (1999): 1 = sem lesão, 2 = 25% de lesão na área foliar, 3 = 50% de lesão na área foliar, 4 = 75% de lesão na área foliar e 5 = área foliar totalmente dessecada. Após o procedimento das análises por meio da abordagem de modelos mistos, os autores observaram variabilidade genética tanto para a característica teor de clorofila quanto para nota de injúria, assim ressaltaram a possibilidade de selecionar genótipos resistentes à *C. oleosa*. Ademais,

evidenciaram correlação positiva entre as características avaliadas, o que permite seleção indireta de genótipos resistentes. Contudo, os autores salientam a preferência pela avaliação por meio do teor de clorofila, argumentando ser uma forma mais fácil e rápida de proceder as avaliações.

Figueiredo et al. (2012) estimaram parâmetros genéticos do valor nutricional e de características agrônômicas de *U. humidicola* utilizando modelos mistos. Com isso propuseram um índice de seleção múltiplo destas características para selecionar genótipos superiores. Dos caracteres considerados, a rebrota, associada a densidade de perfilho das plantas, foram as únicas avaliadas por meio de notas, de modo que a nota 1 determinava uma habilidade de rebrota inferior a 20%, 2 de 20% a 40%, 3 de 40% a 60%, 4 de 60% a 80% e 5 acima de 80%. Dentre as conclusões, os autores relataram que o estudo de várias características para a construção de um índice de seleção múltiplo permite uma melhor compreensão da performance das progênes, assim possibilita uma seleção refinada das plantas mais promissoras por considerar a associação de vários caracteres. Simeão et al. (2016) conduziram um estudo semelhante em *U. ruziziensis* tetraploidizada. Estimaram parâmetros genéticos e fenotípicos, predisseram valores genéticos individuais, estimaram correlações entre características associadas a PMV e valor nutritivo para a comparação de índices de seleção múltipla. Os autores reafirmaram a relevância desse tipo de índice na seleção de plantas de *U. ruziziensis* e no auxílio do surgimento de novas cultivares melhoradas. Além disso, ressaltaram a importância da identificação de parentais de elevada performance produtiva e nutricional, de modo a promover hibridização com outras espécies como *U. brizantha* e *U. decumbens*.

Embora seja evidente as vantagens de se adotar um modelo misto para estimar os efeitos fixos (BLUE) e prever os efeitos aleatórios (BLUP) (HENDERSON, 1959; HENDERSON, 1984; ROBINSON, 1991; PIEPHO, et al., 2008), seu emprego é preconizado às variáveis contínuas, a exemplo da altura de plantas, volume de caldo e produção de grãos. No caso de dados de contagem como estande e florescimento, dados da presença ou ausência de doença ou dados de notas, conforme os citados anteriormente, é indicado a utilização de modelos que atentem para a distribuição da variável resposta (STROUP, 2012, STROUP, 2015), visto que permite análises potencialmente mais poderosas e eficientes.

2.3.2. Transformação Box-Cox

Uma alternativa para análise de dados de notas comparativamente menos utilizada nos dias atuais (JAEGER, 2008, BOLKER, 2009; STROUP, 2015), mas que outrora foi amplamente aplicada e indicada como relevante, é a utilização da transformação na escala da variável, sumarizada por Bartlett (1947). No passado, sua aplicabilidade foi amparada pelo fato de conceder normalidade e homocedasticidade aos erros experimentais, o que permitia acondicionar o processamento de dados por meio das ANAVAs. Porém, atualmente, a heterogeneidade das variâncias já não é um problema eminente, como ressaltado em tópicos anteriores, e a normalidade dos erros é, convenientemente, assegurada pelo teorema do limite central (MOOD; GRAYBILL; BOES, 1974), muito embora seja questionável um número de tratamentos ou repetições que determine a sua viabilidade.

A literatura apresenta aplicações práticas para transformar a escala de dados de notas ressaltando transformações incomuns no melhoramento genético vegetal, a exemplo da transformação arco-seno (BOUTSALIS, 2001) para estudo de resistência à herbicidas e a transformação logit (WANG et al., 2007) para estudo de resistência à pragas. Contudo, o método embasado no conhecimento da relação existente entre a média e a variância dos dados permitiu o desenvolvimento de transformações corriqueiras como a raiz quadrada, a logarítmica e a inversa (BARTLETT, 1936, BARTLETT; KENDALL, 1946), as quais são mais frequentes no contexto de dados de notas em programas de melhoramento de plantas (GRAVINA, et al., 2004; LEMMENS, 2004; LUCIO et al., 2012;). Mesmo assim, em determinados casos, a aplicação dessas transformações não são satisfatórias, sendo necessário a busca por uma transformação mais efetiva. Box e Cox (1964), vislumbrando essas limitações, propuseram uma nova maneira para determinar a transformação na escala da variável por meio de um procedimento analítico que ficou conhecido como transformação Box-Cox.

A transformação Box-Cox foi originalmente apresentada da seguinte forma:

$$y_t(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0 \\ \log y, & \text{se } \lambda = 0 \end{cases} \quad (2.2)$$

em que:

λ : valor do expoente ao qual as notas devem ser elevadas; y : nota; $y_t(\lambda)$: nota transformada.

O parâmetro lambda (λ) pode assumir infinitos valores entre -2 a 2, porém somente aqueles que calibram a transformação, de modo a maximizar a eficiência em mover as notas para a normalidade, que realmente importa. Quando λ assume valor 0, por definição, a transformação se dá pela aplicação do logaritmo natural da própria nota atribuída. Uma forma para se determinar lambda consiste na aplicação da função que maximiza o logaritmo natural da verossimilhança do modelo ajustado para a nota (BOX; COX, 1964).

Alguns casos especiais devem ser ressaltados, já que a transformação Box-Cox incorpora muitas outras transformações tradicionais:

$\lambda = 1$: sem necessidade de transformação; produz resultados idênticos aos dados originais;

$\lambda = 0.50$: transformação raiz quadrada;

$\lambda = 0.33$: transformação raiz cúbica;

$\lambda = 0.25$: transformação raiz a quarta;

$\lambda = -0.50$: transformação recíproca da raiz quadrada;

$\lambda = -1$: transformação recíproca (inversa).

Lúcio et al. (2011) realizaram um estudo com o objetivo de identificar a transformação mais adequada para se avaliar a característica produção de frutos comercializáveis de pimentão em ambiente protegido. Segundo os autores, a previa transformação dos dados foi necessária devido a presença inflacionada de zeros decorrente da ausência de frutos em determinadas parcelas ao longo das colheitas. Assim, poderia causar um impacto negativos nos resultados, caso as análises fossem realizadas com os dados na escala original, em virtude da violação do pressuposto da homocedasticidade da ANAVA. Para isso, recorreram à transformação Box-Cox considerando valores de -2 a 2 para a escolha da melhor potência λ . Concluíram que a transformação mais indicada foi a inversa da raiz quarta por atuar de modo a reduzir a variabilidade e normalizar os resíduos, muito embora não tenha sido eficiente para tornar as variâncias residuais homogêneas.

O avanço dos modelos mistos nas diversas áreas do conhecimento, sobretudo, no melhoramento genético, motivado pelas vantagens da estimação dos componentes da variância, induziram Otsuk et al. (1991) a investigarem a influência da transformação dos dados sobre essas componentes. Para isso, simularam um cenário supondo dados característicos da distribuição Binomial e Poisson, de modo que fosse possível visualizar o efeito da transformação dos dados sobre as componentes de variância. Assim, os autores analisaram a proporção da variação explicada com e sem a transformação dos dados. Segundo o estudo, a proporção da variação em ambos os casos não diferem significativamente, permitindo a estimação das componentes da variância mesmo sobre dados transformados.

Sakia (1990) reportou um viés na utilização da transformação Box-Cox ao apresentar estimadores de máxima verossimilhança de parâmetros para um modelo ANOVA misto balanceado sob dados transformados. Originalmente, os dados descreviam uma escala de notas, empregados, anteriormente, em um estudo da severidade de doença causada por *Cercospora herpotrichodes* (Fron) em progênies de trigo de inverno. Segundo o autor, o viés é constatado quando as análises são reproduzidas na escala original. Comenta que, embora as médias transformadas possam ser consideradas não viesadas, o mesmo não se estende quando apresentadas na escala original. Ademais, o autor ressalta que, quanto mais próximo de zero for o valor de λ , melhor serão as estimativas das variâncias na escala transformada.

Um caso prático da transformação Box-Cox aplicado ao melhoramento genético animal foi executado por Besbes et al. (1991) com o objetivo de analisar a característica produção de ovos em aves e avaliar o efeito da transformação dos dados na seleção dos melhores animais em um programa de melhoramento genético. Para tanto, adotaram um modelo misto geral utilizando o método REML/BLUP para estimação das componentes de variância e predição dos efeitos aleatórios. Segundo os autores, a transformação Box-Cox resultou em um incremento da herdabilidade para o caráter em questão, além de apresentar correlações positivas de elevada magnitude (0,90 a 0,96) entre os valores genéticos estimados dos animais selecionados antes e após as transformações. Assim, os autores concluíram que a adoção da transformação Box-Cox no enfoque de modelos mistos se mostra satisfatória.

A aplicação da transformação Box-Cox, embora se mostre útil pelo fato de solucionar determinados problemas, como os citados anteriormente, produz outros tão críticos quanto os primeiros, possibilitando classificar sua utilização como controversa (STROUP, 2015). Um desses problemas é a perda de um grau de liberdade no resíduo da análise, uma vez que o mesmo conjunto de dados é empregado para estimar o parâmetro λ , o que implica na redução do poder do teste. Outro problema ainda maior é concernente às conclusões acerca dos dados transformados. Enquanto os testes de significância e as comparações de médias devem ser realizados sobre os dados transformados, as inferências e as interpretações práticas devem ser realizadas na escala original, via transformação de médias e intervalos de confiança, de modo a reproduzir a escala original (RESENDE, 2007). Portanto, é razoável pensar que a utilização de modelos mais gerais, que dispensam a necessidade de normalidade de suas variáveis respostas, seja mais indicado para o procedimento das análises estatísticas, visto que estes modelos admitem a escala original dos dados, possibilitando uma interpretação direta dos seus resultados.

2.3.3. Modelos lineares mistos generalizados

Nelder e Wedderburn (1972) desenvolveram os modelos lineares generalizados (MLG) com o objetivo de ampliar as opções para modelar a variável resposta, de modo a reduzir, significativamente, os contrastes relacionados à escala original dos dados experimentais. Assim, apresentaram a distribuição da família exponencial com a finalidade de incluir, em um mesmo grupo, distribuições que compartilhavam propriedade comuns. A distribuição da família exponencial pode ser caracterizada da seguinte forma:

$$f(y_i; \theta_i, \phi) = \exp[\phi\{y_i\theta_i - b(\theta_i)\} + c(y_i, \phi)] \quad (2.3)$$

em que:

y_i : nota atribuída ao genótipo;

θ_i : parâmetros genéticos e ambientais;

ϕ : parâmetro de dispersão;

$b(\cdot)$, $c(\cdot)$: funções específicas.

Uma componente extremamente importante que surge na abordagem dos MLG é a função de ligação $g(\cdot)$ (link function). Sua finalidade consiste em relacionar um preditor linear (η) à esperança dos valores observados $[E(y)]$, logo concede uma maior flexibilidade aos modelos, tornando-os menos rígidos às exigências de normalidade e homogeneidade das variâncias (MCCULLAGH; NELDER, 1989). Em alguns casos, a própria função de ligação classifica o tipo de modelo a ser adotado. Sua forma genérica é apresentada a seguir:

$$g[E(y)] = \eta \quad (2.4)$$

Nos modelos lineares clássicos, a média (μ) e o η são idênticos, logo é plausível admitir uma função de ligação identidade entre eles, já que tanto η quanto μ pode assumir qualquer valor no espaço R . Em contrapartida, quando são abordados dados de contagem assumindo distribuição Poisson, por exemplo, $\mu > 0$, logo a função de ligação identidade é menos atrativa, em partes porque η pode ser negativa enquanto μ não. Segundo McCullagh e Nelder (1989), modelos de contagem, alicerçados na independência de dados cruzados, leva naturalmente a efeitos multiplicativos e se expressam pelo função de ligação \log , $\eta = \log \mu$,

com inversa $\mu = e^\eta$. Desse modo, os efeitos aditivos que contribuem para η , tornam-se efeitos multiplicativos que contribuem para μ , logo μ é necessariamente positivo.

A inclusão de efeitos aleatórios ao MLG promove o surgimento de uma classe de modelos ainda mais abrangente conhecida como modelos lineares mistos generalizados (MLMG) (BRESLOW, CLAYTON, 1993; MCCULLOCH, 1997). Um MLMG é apresentado a seguir:

$$g[E(y|a)] = X\beta + Za \quad (2.5)$$

em que:

$g(\cdot)$: função de ligação

$E[y|a]$: esperança das notas dado os efeitos aleatórios realizados, ($E[y|a] = \mu$)

X : matriz do modelo referente aos efeitos fixos β , de dimensões $n \times p$;

β : vetor dos efeitos fixos, de dimensões $p \times 1$;

Z : matriz do modelo referente aos efeitos aleatórios a , de dimensões $n \times q$;

a : vetor dos efeitos aleatórios, com $a \sim N(0, G)$;

Desse modo, fica explícito que os modelos citados até o momento são casos particulares dos MLMG, em que certas restrições são admitidas.

A flexibilização concedida pela abordagem de MLMG demanda um conhecimento mais detalhado a respeito da variável utilizada no estudo da característica, bem como da distribuição que a descreve (BOLKER, 2009). No que se refere a dados de notas para estudo da característica vigor da planta, a variável nota assume a função de classificação dessas plantas, de modo que as menores notas sejam atribuídas àquelas de menor performance e, consequentemente, aquelas de melhor performance recebam as maiores notas. Por essa razão é razoável classificar essa variável como categórica ordinal (AGRESTI, 2002). Nesse caso, o interesse do pesquisador inclui o conhecimento das probabilidades de encontrar genótipos presentes nas categorias superiores, logo exige a adoção de uma distribuição que consiga estimar essas probabilidades. Segundo a literatura, a distribuição mais indicada para efetuar essa estimação é a multinomial (JANSEN, 1991; AGRESTI, 2002; AGRESTI, 2007), comumente apresentada da seguinte forma:

$$p(n_1, n_2, \dots, n_c) = \left(\frac{n!}{n_1! n_2! \dots n_c!} \right) \pi_1^{n_1} \pi_2^{n_2} \dots \pi_c^{n_c} \quad (2.6)$$

em que:

c : número de categorias;

π : probabilidades de cada categoria, sendo $\sum_j \pi_j = 1$;

n_i : frequência da nota relacionada a categoria i .

O procedimento para análise de dados que seguem uma distribuição multinomial é comumente realizado por meio dos modelos thresholds (ou de limiar) (FALCONER; MACKAY, 1996) sob aplicação da função de ligação probit ou logit, (acumulado, adjacente ou de referência) dependendo da finalidade do experimento (AGRESTI, 2002; STROUP, 2012). À esses modelos, compete a estimação de limiares (*cut-points*) responsáveis por delimitar categorias, logo possibilita, por exemplo, a classificação de genótipos conforme a categoria à qual ele esteja inserido. É importante ressaltar que o posicionamento de um genótipo à uma determinada categoria é consequência do mérito genético desse genótipo (FALCONER; MACKAY, 1996). No contexto do melhoramento genético, cuja finalidade é a seleção de genótipos superiores, os modelos mais indicados são os probit ou logit acumulado (AITCHISON, SILVEY, 1957; TEMPELMAN, 1998; MIELENZ, 2010; MONTESINOS-LÓPEZ, 2015), já que permitem, de modo prático, a discriminação de categoria superiores. Devido a adoção do modelo threshold com aplicação da função de ligação probit para procedimento das análises desse trabalho, somente estes serão apresentados.

A função probit (Φ) é definida como sendo a função densidade cumulada normal de variação 0 a 1 e pode ser aplicada diretamente em modelos não lineares com a finalidade de linearizá-los por meio da expressão $\pi(n) = \Phi(X\beta + Za)$. Quando Φ é estritamente crescente no espaço R , sua inversa ($\Phi^{-1}(\cdot)$) existe e é denominada de função de ligação probit nos MLMG (AGRESTI, 2002). Assim um modelo probit na abordagem de MLMG pode ser apresentado da seguinte maneira:

$$\Phi^{-1}(\pi_c | a) = X\beta + Za \quad (2.7)$$

em que:

Φ^{-1} : função de ligação probit;

π_c : probabilidade das categorias, sendo c o índice indicativo da categoria;

X : matriz do modelo referente aos efeitos fixos β , de dimensões $n \times p$;

β : vetor dos efeitos fixos, de dimensões $p \times 1$;

Z : matriz do modelo referente aos efeitos aleatórios u , de dimensões $n \times q$;

a : vetor dos efeitos aleatórios, com $a \sim N(0, G)$;

Conforme já mencionado, modelos threshold fornecem um meio para modelagem de dados categóricos ordinais, porém a interpretação de seus resultados difere da dos modelos cujos dados são normalmente distribuídos. Isso decorre do fato de que os modelos que pressupõe normalidade, invariavelmente pressupõe também aditividade de seus efeitos, isto é, a mudança no valor da variável dependente implica no aumento ou diminuição da média do tratamento. Em contrapartida, para os modelos que consideram a natureza dos dados categóricos, tal assunção é impropriedade, visto que seus efeitos são caracterizados pelos impactos sobre a probabilidade de um tratamento cair numa determinada categoria. Nos modelos threshold com aplicação da função de ligação probit, por exemplo, a menos que um genótipo se encontre acima de um certo limiar, ele será observado na categoria inferior. Stroup (2012) distingue os dois modelos citados em modelos de equação (distribuição normal) e modelos de probabilidade (distribuição multinomial).

Os modelos threshold são amplamente abordados no contexto do melhoramento genético animal (GIANOLA, 1981; GIANOLA, 1982; MIELENZ, 2010; TEMPELMAN, 1998), em que sua aplicação é considerada eficiente para selecionar animais indiretamente por meio dos escores visuais. Contudo, essa realidade não se estende ao cenário do melhoramento vegetal, possibilitando o questionamento da sua eficiência nesse caso.

2.3.1. Modelos lineares mistos generalizados bayesianos

Os modelos apresentados anteriormente contemplam a filosofia frequentista de estimação de parâmetros. Por muito tempo, ela foi considerada a única forma viável para execução de análises estatísticas. Contudo, a revolução no campo da computação e da tecnologia permitiu a retomada da filosofia bayesiana, visto que seus estimadores já não eram considerados impraticáveis. Os modelos frequentistas, segundo alguns pesquisadores, assumem muitas aproximações e forte suposições que facilmente são violadas, podendo gerar estimativas e previsões espúrias dos experimentos. Em face dessas limitações, os modelos

bayesianos têm sido adotados no melhoramento genético vegetal e, principalmente, animal, como uma ferramenta poderosa para resolução de problemas relacionados à estimação de parâmetros (GIANOLA, FOULLEY, 1983; GIANOLA, FERNANDO, 1986).

A inferência bayesiana teve sua origem em 1763, quando Richard Price publicou a obra póstuma de Thomas Bayes intitulada “An Essay Towards Solving a Problem in the Doctrine of Chances”. Uma das vantagens frequentemente ressaltadas por bayesianos é o fato dessa estatística ser alicerçada em apenas uma ferramenta (teorema de Bayes) para revisar suas convicções acerca dos parâmetros frente aos dados experimentais, enquanto os modelos frequentistas adotam mais de uma ferramenta para fazer o mesmo. Outra grande vantagem dos modelos bayesianos é a realização das inferências sobre as distribuições de probabilidade dos parâmetros, permitindo a construção de intervalos de credibilidade, enquanto que os frequentistas realizam inferências sobre os estimadores dos parâmetros, viabilizando a construção dos intervalos de confiança (BOLSTAD, 2013).

O teorema de Bayes pode ser definido da seguinte maneira:

$$P(\theta | y) \propto P(y | \theta)P(\theta) \quad (2.8)$$

em que:

$P(\theta | y)$: distribuição *a posteriori* de θ ;

$P(y | \theta)$: distribuição conjunta dos dados, dado os parâmetros (verossimilhança);

$P(\theta)$: distribuição *a priori* de θ .

Dessa forma, fica explícito que a probabilidade *a posteriori* da ocorrência de um evento é proporcional a verossimilhança da ocorrência desse evento ponderado por um conhecimento *a priori*.

A inserção da informação *a priori* nos modelos bayesianos permite a obtenção de estimativas mais acuradas dos parâmetros (BOLSTAD, 2013). A informação *a priori* é a distribuição assumida para o parâmetro, independentemente dos dados. É por meio desta distribuição que o pesquisador incorpora o conhecimento prévio a respeito do parâmetro a ser estimado. Uma maneira de se obter uma distribuição *a priori* é por meio da meta-análise, a qual consiste em levantar estudos passados que propuseram a estimação de parâmetros, de modo a identificar o comportamento destes. Em estudos prospectivos, a distribuição *a priori* pode ser usada como um mecanismo de atualização de informação, em que a informação *a*

posteriori torna-se a informação *a priori* em estudos futuros. No melhoramento genético, uma informação *a priori* muito importante é a herdabilidade do caráter (SORENSEN, GIANOLA, 2002).

Embora seja coerente pensar que a inclusão de um conhecimento prévio acerca dos parâmetros a serem estimados é uma vantagem, o emprego das prioris não é um consenso na comunidade científica. Segundo Sorensen e Gianola (2002), uma das maiores críticas à inferência Bayesiana é o potencial efeito que uma priori subjetiva, arbitrária e/ou equivocada tem sobre as inferências. Assim, estudos com essa abordagem têm sido direcionados pela adoção de prioris não informativas, aquelas cujas distribuições contribuem pouco nas estimativas paramétricas (JEFFREYS, 1961; BOX, TIAO, 1973). Alguns desses estudos comprovam que o efeito de uma priori é dissipado com o aumento do tamanho da amostra. Contudo, quando as observações da amostra é correlacionada ou quando o modelo envolve muitos parâmetros, nem sempre é claro o quão grande deve ser essa amostra para que a influência da priori seja encoberta pelos dados (SORENSEN, GIANOLA, 2002).

Mesmo não havendo consenso sobre a utilização de prioris informativa ou não informativa, é importante ressaltar que o emprego da inferência Bayesiana na análise de experimentos tem se intensificado com o passar dos anos (GOUY et al., 2013; MILA, NGUGI, 2011). Um exemplo claro do avanço da Bayesiana é o desenvolvimento de pacotes estatísticos específicos para análise dados categóricos, como é o caso do pacote Bayesthresh (CORREA, 2016), o que, invariavelmente, estimula pesquisadores a considerar essa abordagem para processamento de dados de notas. Na apresentação do pacote supracitado, o autor propõe um modelo threshold supondo uma priori gama inversa para as componentes das variâncias dos efeitos, indicando ser uma alternativa comum para estimativa de parâmetros dessa natureza. Assim, ilustra a aplicação do pacote por meio da análise de um experimento com variedades de tomateiro visando a seleção para a característica resistência à requeima causada pelo fungo *Phytophthora infestans*, evidenciando a intensificação da aplicação do emprego da Bayesiana inclusive no melhoramento genético de plantas.

Os modelos threshold para análise de dados categóricos podem ser utilizados tanto na filosofia frequentista quanto na bayesiana. (GIANOLA, NORTON, 1981; GIANOLA, FOULLEY, 1983; SORENSEN et al. 1995; LUO, 2001; GIANOLA, 2013). Entretanto, a maioria dos estudos práticos que utilizam esses modelos estão relacionados ao melhoramento genético animal. Portanto é necessário que haja uma exploração maior dos modelos threshold no enfoque do melhoramento vegetal, para que se estabeleça as vantagens e desvantagens da sua aplicação, de modo que seja esclarecido a efetividade ou não da sua adoção.

A definição de um modelo linear misto generalizado bayesiano (MLMGB) é feita conforme apresentada em 2.7. Portanto, é possível afirmar que os MLMG e MLMGB podem especificar um mesmo modelo da mesma forma, sendo a inclusão da informação *a priori* nos MLMGB a única distinção nessa especificação. Vale ressaltar que, quando não se dispõe de uma priori informativa, as estimativas dos MLMG não diferem das estimativas dos MLMGB (SORENSEN, GIANOLA, 2002). No presente trabalho, a análise dos dados de notas empregou o modelo threshold assumindo a priori gama inversa para a distribuição das componentes de variância, seguindo a indicação da literatura.

3. CONSIDERAÇÕES GERAIS

A intenção deste trabalho é proporcionar ao leitor um maior conhecimento a respeito da *Urochloa ruziziensis*, mostrar a importância desta espécie no processo do melhoramento genético de forrageiras e reforçar a utilidade da avaliação visual nesse contexto. Além disso, objetiva facilitar a compreensão de abordagens estatísticas para análise de notas que são pouco contemplados no melhoramento de plantas, expondo conceitos, por meio de um estudo prático, e apresentando as justificativas relevantes que conduzem a opção pelo emprego de uma determinada abordagem.

REFERÊNCIAS

- ABADASSI, J., Mass Selection to Improve Grain Yield in a Traditional Tropical Maize Population. **International Journal of Science and Technology**, v. 4, n. 3, 2014.
- ABREU, G. B.; RAMALHO, M. A. P.; TOLEDO, F. H. R. B.; SOUZA, J. C. de. Strategies to improve mass selection in maize. **Maydica**, Bergamo, v. 55, n. 4, p. 219-225, 2010.
- AGRESTI, A. **An Introduction to Categorical Data Analysis** (second ed.), News York: John Wiley & Sons. 2007. 372p.
- AGRESTI, A. **Categorical data analysis** (second ed.). New York, NY: John Wiley & Sons, 2002. 710p.
- AGUIRRE, L. M.; CARDONA, C.; MILES, J. W.; SOTELO, G. Characterization of resistance to adult spittlebugs (Hemiptera: Cercopidae) in *Brachiaria* spp. **Journal of economic entomology**, v. 106, n. 4, p. 1871-1877, 2013.
- AITCHISON, J.; SILVEY, S. D. The generalization of probit analysis to the case of multiple responses. **Biometrika**, v. 44, n. 1/2, p. 131-140, 1957.
- ATKINS, R. E. Visual Selection for Grain Yield in Barley. **Crop Science**, v. 4, n. 5, p. 494-497, 1964.
- BARTLETT, M. S. The use of transformations. **Biometrics**, v. 3, n. 1, p. 39-52, 1947.
- BARTLETT, M. S.; KENDALL, D. G. The statistical analysis of variance-heterogeneity and the logarithmic transformation. **Supplement to the Journal of the Royal Statistical Society**, v. 8, n. 1, p. 128-138, 1946.
- BARTLETT, M.S. The square root transformation in analysis of variance. **Supplement to the Journal of the Royal Statistical Society**, v. 3, n. 1, p. 68-78, 1936.
- BAYES, T.; PRICE, R.; CANTON, J. **An essay towards solving a problem in the doctrine of chances**. C. Davis, Printer to the Royal Society of London, 1763.
- BERNARDO, R. **Breeding for quantitative traits in plants**. 2^a Ed. Woodbury: Stemma Press, 2010. 400 p.
- BESBES, B. et al. Box-Cox transformation of egg-production traits of laying hens to improve genetic parameter estimation and breeding evaluation. **Livestock Production Science**, v. 33, n. 3-4, p. 313-326, 1993.
- BOLKER, B. M. et al. Generalized linear mixed models: a practical guide for ecology and evolution. **Trends in ecology & evolution**, v. 24, n. 3, p. 127-135, 2009.
- BOLSTAD, W. M. **Introduction to Bayesian statistics**. John Wiley & Sons, 2013.
- BOUTSALIS, P. Syngenta Quick-Test: A Rapid Whole-Plant Test for Herbicide Resistance 1. **Weed Technology**, v. 15, n. 2, p. 257-263, 2001.
- BOX, G. E. P.; TIAO G. C. **Bayesian Inference in Statistical Analysis**. Wiley. 1973.
- BOX, G. E. P.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 211-252, 1964.

- BOX, G. E. P.; DRAPER N. R. **Empirical model-building and response surfaces**. Vol. 424. New York: Wiley, 1987.
- BRESLOW, N. E.; CLAYTON, D. G. Approximate inference in generalized linear mixed models. **Journal of the American statistical Association**, v. 88, n. 421, p. 9-25, 1993.
- BRUCKNER, P. L.; RAYMER, P. L.; BURTON, G. W. Recurrent phenotypic selection for forage yield in rye. **Euphytica**, v. 54, n. 1, p. 11-17, 1991.
- BURTON, G. W. Improved recurrent restricted phenotypic selection increases bahiagrass forage yields. **Crop Science**, v. 22, n. 5, p. 1058-1061, 1982.
- BURTON, G. W. Recurrent restricted phenotypic selection increases forage yields of *Pensacola bahiagrass*. **Crop Science**, v. 14, n. 6, p. 831-835, 1974.
- CARDONA, C.; FORY, P.; SOTELO, G.; PABÓN, A.; DIAZ, G.; MILES, J. W. Antibiosis and tolerance to five species of spittlebug (Homoptera: Cercopidae) in *Brachiaria spp.*: implications for breeding for resistance. **Journal of economic entomology**, v. 97, n. 2, p. 635-645, 2004.
- CARDONA, C.; MILES, J. W.; SOTELO, G. An improved methodology for massive screening of *Brachiaria spp.* genotypes for resistance to *Aeneolamia varia* (Homoptera: Cercopidae). **Journal of economic entomology**, v. 92, n. 2, p. 490-496, 1999.
- CARDONA, C.; MILES, J. W.; ZUÑIGA, E.; SOTELO, G. Independence of resistance in *Brachiaria spp.* to nymphs or to adult spittlebugs (Hemiptera: Cercopidae): implications for breeding for resistance. **Journal of economic entomology**, v. 103, n. 5, p. 1860-1865, 2010.
- CASLER, M. D.; VOGEL, K. P. Selection for biomass yield in upland, lowland, and hybrid switchgrass. **Crop Science**, v. 54, n. 2, p. 626-636, 2014.
- CORRÊA, F. M. et al. Bayesian algorithms for analysis of categorical ordinal data. **Revista brasileira de biometria/biometric brazilian journal**, v. 34, n. 4, p. 597-620, 2016.
- CROSSA, J. et al. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. **Genetics**, v. 186, n. 2, p. 713-724, 2010.
- CUTRIM, V.; RAMALHO, M. A. P.; CARVALHO, A. M. Eficiência da seleção visual na produtividade de grãos de arroz (*Oryza sativa* L.) irrigado. **Pesquisa Agropecuária Brasileira**, v. 32, n. 6, p. 601-606, 1997.
- EISENHART, C. The assumptions underlying the analysis of variance. **Biometrics**, Raleigh, v. 3, n. 1, p. 1-21, 1947.
- FALCONER, D. S.; MACKAY, T. F. C. Threshold characters. **Introduction to Quantitative genetics**, p. 299-311, 1996.
- FÈ, D. et al. Genetic and environmental variation in a commercial breeding program of perennial ryegrass. **Crop Science**, v. 55, n. 2, p. 631-640, 2015.
- FERGUSON, J. E.; CROWDER, L. V. Cytology and breeding behavior of *Brachiaria ruziziensis* Germain et Evrard. **Crop Science**. Madison, v. 14, p. 893-895, 1974.

- FIGUEIREDO, U. J.; NUNES, J. A. Rodrigues; VALLE, C. B. do. Estimation of genetic parameters and selection of *Brachiaria humidicola* progenies using a selection index. **Crop Breeding and Applied Biotechnology**, v. 12, n. 4, p. 237-244, 2012.
- GIANOLA, D. Theory and analysis of threshold characters. **Journal of Animal Science**, v. 54, n. 5, p. 1079-1096, 1982.
- GIANOLA, D.; FERNANDO, R.L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v.63, p.217-244, 1986.
- GIANOLA, D.; FOULLEY, J. L. Sire evaluation for ordered categorical data with a threshold model. **Genetics Selection Evolution**, v. 15, n. 2, p. 1-23, 1983.
- GIANOLA, D.; NORTON, H. W. Scaling threshold characters. **Genetics**, v. 99, n. 2, p. 357-364, 1981.
- GOBBE, J.; SWENNE, A.; LOUANT, B. P. Diploïdes naturels et autotétraploïdes induits chez *Brachiaria ruziziensis* Germain et Evrard: critères d'identification. **Agronomy Tropical**, v.36, p.339-346, 1981.
- GOUY, M. et al. Experimental assessment of the accuracy of genomic selection in sugarcane. **Theoretical and Applied Genetics**, v. 126, n. 10, p. 2575-2586, 2013.
- GRAVINA, G. D. A., CECON, P. R., MARTINS FILHO, S., SEDIYAMA, C. S. Transformação de dados nas estimativas de parâmetros genéticos da resistência dos cultivares de soja Uberaba e Bossier à *Cercospora sojina* Hara. **Ceres**, 2004.
- HACKER J. B., JANK L. Breeding tropical and subtropical forage plants. In: Cherney, J.H. & D.J.R. Cherney (Eds.) **Grass for dairy cattle**. Oxfordshire, CABI Wallingford. p. 49-71. 1998.
- HADFIELD, J. D. et al. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. **Journal of Statistical Software**, v. 33, n. 2, p. 1-22, 2010.
- HADFIELD, J. D. MCMCglmm course notes. **Unpublished Manuscript, University of Edinburgh**, 2012.
- HALLAUER, A. R., Evolution of plant breeding. **Crop Breeding and Applied Biotechnology**, v. 11, n. 3, p. 197-206, 2011.
- HENDERSON, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, p. 423-447, 1975.
- HENDERSON, C R. et al. The estimation of environmental and genetic trends from records subject to culling. **Biometrics**, v. 15, n. 2, p. 192-218, 1959.
- HENDERSON, C. R. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. **Biometrics**, p. 69-83, 1976.
- HENDERSON, C. R. **Applications of linear models in animal breeding**. 384 p. 1984.
- JAEGER, T. F. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. **Journal of memory and language**, v. 59, n. 4, p. 434-446, 2008.

- JANK L., CARVALHO P. De F., VALLE C.B. do. New grasses and legumes: advances and perspectives for the tropical zones of Latin America. In: Reynolds SG & Frame J (Org.) Grasslands: developments, opportunities, perspectives. Roma, **FAO**, India, Science Publishers. p. 55-79. 2005.
- JANK, L.; BARRIOS, S. C.; VALLE, C. B. do.; SIMEÃO, R. M.; ALVES, G. F. The value of improved pastures to Brazilian beef production. **Crop & Pasture Science**. **Clayton**, v. 65, n. 11, p. 1132-1137, out. 2014.
- JANSEN, J. Fitting regression models to ordinal data. **Biometrical Journal**, v. 33, n. 7, p. 807-815, 1991.
- JEFFREYS, H. **Theory of Probability**. Clarendon Press. 1961.
- LEMMENS, M.; HAIM, K.; LEW, H.; RUCKENBAUER, P. The effect of nitrogen fertilization on *Fusarium* head blight development and deoxynivalenol contamination in wheat. **Journal of Phytopathology**, v. 152, n. 1, p. 1-8, 2004.
- LOPES, F. C. et al. Composição química e digestibilidade ruminal in situ da forragem de quatro espécies do gênero *Brachiaria*. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 62, n. 4, p. 883-888, 2010.
- LÓPEZ, F.; CARDONA, C.; MILES, J. W.; SOTELO, G.; MONTOYA, J. Screening for resistance to adult spittlebugs (Hemiptera: Cercopidae) in *Brachiaria spp.*: methods and categories of resistance. **Journal of economic entomology**, v. 102, n. 3, p. 1309-1316, 2009.
- LÚCIO, A. D. C., COUTO, M. R. M., LOPES, S. J., STORCK, L. Transformação box-cox em experimentos com pimentão em ambiente protegido. **Horticultura Brasileira**, 2011.
- LÚCIO, A. D., SCHWERTNER, D. V., HAESBAERT, F. M., SANTOS, D., BRUNES, R. R., RIBEIRO, A. L., LOPES, S. J. Violação dos pressupostos do modelo matemático e transformação de dados. **Horticultura Brasileira**, 2012.
- LUO, M. F. et al. Bayesian inference for categorical traits with an application to variance component estimation. **Journal of Dairy Science**, v. 84, n. 3, p. 694-704, 2001.
- MAASS, B. L., Identifying and Naming *Brachiaria* Species In: MILES, J.W.; MAASS, B.L.; VALLE, C.B. do, ed. **Brachiaria: Biology, Agronomy, and Improvement**. Cali: CIAT/Brasília: EMBRAPA-CNPGC, 1996.
- MCCULLAGH, P.; NELDER, James A. **Generalized Linear Models**, no. 37 in Monograph on Statistics and Applied Probability. 1989.
- MCCULLOCH, C. E. et al. **An Introduction to Generalized Linear Mixed Models**. Biometrics Unit Technical Reports: Number BU-1340-MA. 1997.
- MCCULLOCH, C. E.; SEARLE, S. R.; NEUHAUS, J. M. **Generalized, Linear, and Mixed Models**. John Wiley & Sons, 2008.
- MIELENZ, N.; SPILKE, J.; VON BORELL, E. Analysis of ordered categorical data with threshold models exemplified by plumage damage scores from laying hens differing in their genotype and rearing environment. **Poultry science**, v. 89, n. 11, p. 2521-2534, 2010.
- MIGUEL, P. S. B. et al. Seleção de genótipos de *Brachiaria Ruziziensis* quanto ao alumínio em solução nutritiva: II: Avaliação da tolerância ao alumínio. **Revista de Ciências Agrárias**, v. 34, n. 1, p. 163-172, 2011.

- MILA, A. L.; NGUGI, H. K. A Bayesian approach to meta-analysis of plant pathology studies. **Phytopathology**, v. 101, n. 1, p. 42-51, 2011.
- MILES J.W., VALLE C. B. Do. Manipulation of apomixis in *Brachiaria* breeding In: MILES JW, MAASS BL & VALLE CB do (Eds.) **Brachiaria: biology, agronomy, and improvement**. CIAT/Brasília: EMBRAPA-CNPQC, 1996.
- MILES, J. W.; CARDONA, C.; SOTELO, G. Recurrent Selection in a Synthetic *Brachiariagrass* Population Improves Resistance to Three Spittlebug Species. **Crop Science**. Madison, v. 46, mar. 2006.
- MONTESINOS-LÓPEZ, O. A. et al. Threshold models for genome-enabled prediction of ordinal categorical traits in plant breeding. **G3: Genes|Genomes|Genetics**, v. 5, n. 2, p. 291-300, 2015.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**, v. 3. 564 p. 1974.
- NANDRAM, B.; CHEN, M. Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. **Journal of Statistical Computation and Simulation**, v. 54, n. 1-3, p. 129-144, 1996.
- NDIKUMANA, J. Etude de l'hybridation entre espèces apomictiques et sexuées dans le genre *Brachiaria*. **Ph.D Thesis**. Université Catholique de Louvain, Louvain-la-Neuve, Belgique. 210 p. 1985.
- NGUYEN, H. T.; SLEPER, D. A. Genetic variability of seed yield and reproductive characters in tall fescue. **Crop science**, v. 23, n. 4, p. 621-626, 1983.
- Nelder, J. A., Wedderburn, R. W. M. Generalized Linear Models. **Journal of the Royal Statistical Society. Series A (General)**. Vol. 135, No. 3, pp. 370-384. (1972)
- NUNES, J. A. R. Incorporação da informação de parentesco no método genealógico pelo enfoque de modelos mistos. **Tese de Doutorado** 2006.
- NUNES, J. A. R.; RAMALHO, M. A. P.; FERREIRA, D. F. Inclusion of genetic relationship information in the pedigree selection method using mixed models. **Genetics and Molecular Biology**, v. 31, n. 1, p. 73-78, 2008.
- OTSUK, I. P., BARBIN, D., & CORRENTE, J. E. Influência de transformações de dados sobre a estimação dos componentes de variância. **Boletim de Indústria Animal**, 1991.
- PABÓN, A.; CARDONA, C.; MILES, J. W.; & SOTELO, G. Response of resistant and susceptible *Brachiaria spp.* genotypes to simultaneous infestation with multiple species of spittlebugs (Hemiptera: Cercopidae). **Journal of economic entomology**, v. 100, n. 6, p. 1896-1903, 2007.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v.58, p.545-554, 1971.
- PEREIRA, A. V. et al. Selection of interspecific *Brachiaria* hybrids to intensify milk production on pastures. **Crop Breeding and Applied Biotechnology**, v. 5, n. 1, p. 99-104, 2005.
- PIEPHO, H. P., MÖHRING, J., MELCHINGER, A. E., & BÜCHSE, A. BLUP for phenotypic selection in plant breeding and variety testing. **Euphytica**, 2008.
- PRICE, D. L.; CASLER, M. D. Divergent selection for secondary traits in upland tetraploid switchgrass and effects on sward biomass yield. **BioEnergy Research**. 2014.

RENCHEER, A. C.; SCHAALJE, G. B. **Linear models in statistics**. John Wiley & Sons, 2008.

RENVOIZE, S.A.; CLAYTON, W.D.; KABUYE, C.H.S. Morphology, Taxonomy, and Natural Distribution of *Brachiaria* (Trin.) Griseb. In: MILES, J.W.; MAASS, B.L.; VALLE, C.B. do, ed. **Brachiaria: Biology, Agronomy, and Improvement**. Cali: CIAT/Brasília: EMBRAPA-CNPQC, 1996.

RESENDE, M. D. V. **Análise estatística de modelos mistos via REML/BLUP na experimentação em melhoramento de plantas perenes**. Embrapa Florestas, 2000.

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

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

RIDAY, H. Correlations between visual biomass scores and forage yield in space planted red clover (*Trifolium pratense L.*) breeding nurseries. **Euphytica**, v. 170, n. 3, p. 339-345, 2009.

ROBINSON, G. K. That BLUP is a good thing: the estimation of random effects. **Statistical science**, p. 15-32, 1991.

SAKIA, R. M. Retransformation bias: a look at the Box-Cox transformation to linear balanced mixed ANOVA models. **Metrika**, v. 37, n. 1, p. 345-351, 1990.

SAVIDAN, Y. H., JANK, L., COSTA, J. C. G.; VALLE, C. D. Breeding *Panicum maximum* in Brazil. 1. Genetic resources, modes of reproduction and breeding procedures. **Euphytica**, v. 41, n. 1, p. 107-112, 1989.

SEARLE, S. R. **Linear Models**. John Wiley & Sons, 1971.

SEARLE, S. R.; CASELLA, G.; MCCULLOCH, C. E. **Variance components**. John Wiley & Sons, 1992.

SERRÃO, E. A.; SIMÃO NETO, M. Informações sobre duas espécies de gramíneas forrageiras do gênero *Brachiaria* na Amazônia: *Brachiaria decumbens* Stapf e *B. ruziziensis* Germain e Everaed. **Boletim do Instituto de Pesquisa Experimental Agropecuária do Norte**, Belém, v. 1, n. 1, p. 1-31, 1971.

SILVA, D. M. et al. Genetic Variability of *Brachiaria ruziziensis* Clones to *Collaria oleosa* (Hemiptera: Miridae) Based on Leaf Injuries. **American Journal of Plant Sciences**, v. 4, n. 12, p. 2418, 2013.

SIMEÃO, R. et al. Genetic evaluation and selection index in tetraploid *Brachiaria ruziziensis*. **Plant Breeding**, 2016.

SMITH, K. F.; TASNEEM, M.; KEARNEY, G. A.; REED, K. F. M.; LEONFORTE, A. Evaluation of herbage yield in a forage grass breeding program: comparison of visual rating versus measurement in single-row plots or swards. **Animal Production Science**, v. 41, n. 8, p. 1161-1166, 2001.

SMITH, L. H. The effect of selection upon certain physical characters in the corn plant. **Illinois Agricultural Experiment Station**, Urbana, n. 132, p. 48-62, fev. 1909.

- SORENSEN, D. A. et al. Bayesian inference in threshold models using Gibbs sampling. **Genetics Selection Evolution**, v. 27, n. 3, p. 229-249, 1995.
- SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian and MCMC Methods in Quantitative Genetics**. Springer, 2002.
- SOUZA SOBRINHO, F. de.; AUAD, A. M.; SILVA LÉDO, F. J. da Genetic variability in *Brachiaria ruziziensis* for resistance to spittlebugs. **Crop Breeding & Applied Biotechnology**, v. 10, n. 1, 2010b.
- SOUZA SOBRINHO, F. D., PEREIRA, A. V., LEDO, F. D. S., BOTREL, M. A., OLIVEIRA, J. S., XAVIER, D. F. Avaliação agronômica de híbridos interespecíficos entre capim-elefante e milho. **Pesquisa Agropecuária Brasileira**, v. 40, n. 9, p. 873-880, 2005.
- SOUZA SOBRINHO, F. de et al. Repetibilidade de características agronômicas e número de cortes necessários para seleção de *Urochloa ruziziensis*. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 6, p. 579-584, jun. 2010a.
- SOUZA SOBRINHO, F.; LEDO, F. J. S.; KOPP, M. M.; PEREIRA, A. V.; SOUZA, F. F. Melhoramento de gramíneas forrageiras na Embrapa Gado de Leite. In: SIMPÓSIO E CONGRESSO DE FORRAGICULTURA E PASTAGENS, 7., 2009, Lavras. **Anais...** Lavras: Universidade Federal de Lavras. p. 98-115, 2009.
- STROUP, W. W. **Generalized linear mixed models: modern concepts, methods and applications**. CRC press, 2012.
- STROUP, W. W. Rethinking the analysis of non-normal data in plant and soil science. **Agronomy Journal**, v. 107, n. 2, p. 811-827, 2015.
- STUTHMAN, D. D.; STEIDL, R. P. Observed gain from visual selection for yield in diverse oat populations. **Crop Science**, v. 16, n. 2, p. 262-264, 1976.
- SWENNE, A. et al. Induction par l'accolchicine de formes autotétraploïdes chez *Brachiaria ruziziensis* Germain et Evrard (Graminée). **AgronomieTropicale**, v.36, n.2, p.134-141, 1981.
- TEIXEIRA, D. H. L. Eficiência da seleção visual para produtividade de biomassa verde de *Brachiaria ruziziensis*. **Tese de doutorado**. Apresentado à Universidade Federal de Lavras – UFLA. 2015.
- TEMPELMAN, R. J. Generalized Linear Mixed Models in Dairy Cattle Breeding. **Journal of dairy science**, v. 81, n. 5, p. 1428-1444, 1998.
- TIMBO, A. L. de O. et al. Obtaining tetraploid plants of ruzigrass (*Brachiaria ruziziensis*). **R. Bras. Zootec.**, Viçosa, v. 43, n. 3, p. 127-131, Mar. 2014.
- VALLE, C. B. do; MACEDO, M. C. M.; EUCLIDES, V. P. B.; JANK, L.; RESENDE, R. M. S. Gênero *Brachiaria*. In: FONSECA, D. M.; MARTUSCELLO, J. A., **Plantas Forrageiras**. Ed. UFV p. 30-77. 2013.
- VALLE, C. B.; JANK, L.; RESENDE, R. M. S. O melhoramento de forrageiras tropicais no Brasil. **Revista Ceres**, v. 56, n. 4, p. 460-472, 2009.
- VALLE, C.B. do; EUCLIDES, V. P. B.; MACEDO, M. C. M. Características das plantas forrageiras do gênero *Brachiaria*. SIMPÓSIO SOBRE MANEJO DA PASTAGEM, v. 17, p. 65-108, 2000.
- WANG, R.; RIPLEY, V. L.; RAKOW, G. Pod shatter resistance evaluation in cultivars and breeding lines of *Brassica napus*, *B. juncea* and *Sinapis alba*. **Plant Breeding**, v. 126, n. 6, p. 588-595, 2007.

WILKINSON, G. N.; ROGERS, C. E. Symbolic description of factorial models for analysis of variance. **Applied Statistics**, p. 392-399, 1973.

WILKINSON, J. M.; HILL, J. Effect on yield and dry-matter distribution of the stay-green characteristic in cultivars of forage maize grown in England. **Grass and Forage Science**, v. 58, n. 3, p. 258-264, 2003.

SEGUNDA PARTE – ARTIGOS

ARTIGO 1

NÚMERO DE AVALIADORES NA INSPEÇÃO VISUAL DA PRODUÇÃO DE MASSA VERDE EM *Urochloa ruziziensis*

RESUMO

A avaliação visual é uma técnica frequentemente utilizada em fases avaliativas dos ciclos de seleção de genótipos de *Urochloa ruziziensis* (sin. *Brachiaria ruziziensis*) para variadas características. O número adequado de avaliadores é um dos questionamentos recorrente relacionados a eficiência da sua utilização. Este trabalho teve por objetivo verificar as implicações do número de avaliadores na eficiência da avaliação visual para vigor de plantas em *U. ruziziensis* e determinar o número otimizado de avaliadores. Para isso foram avaliados 2204 genótipos de *U. ruziziensis* no delineamento em blocos aumentados utilizando duas testemunhas: cultivares Marandu (*U. brizantha*) e Basilisk (*U. decumbens*). Foi mensurada a produção de massa verde (PMV) e atribuído notas de vigor às plantas pela inspeção visual efetuada por seis avaliadores. Os dados foram analisados empregando a abordagem de modelos mistos com recuperação de informação interblocos e intergenotípica. Para o caso das notas, procedeu-se as análises considerando a nota de cada avaliador e as médias das combinações de 2, 3, 4, 5 e 6 avaliadores. Foram estimadas as componentes de variância, a herdabilidade, a acurácia seletiva e o coeficiente de variação experimental (C_{Ve}). A variância genética foi significativa para ambas as características. Constatou-se que a seleção indireta de genótipos produtivos por meio das notas foi eficiente. A correlação entre PMV e nota de vigor para seis avaliadores foi elevada (0,86). A quantidade de avaliadores empregados teve repercussão sobre as estimativas da acurácia seletiva, variância genética e C_{Ve}, sendo desfavorável quando utilizado um único avaliador. Com base nos resultados, é possível afirmar que o número otimizado de avaliadores para fins de seleção de genótipos de *U. ruziziensis* para produção de biomassa é três.

Palavras-chave: Avaliação visual, Acurácia seletiva, Correlação, Melhoramento de forrageiras.

ABSTRACT

Visual evaluation is a technique frequently used in evaluation stages of *Urochloa ruziziensis* (syn. *Brachiaria ruziziensis*) genotype selection cycles for a variety of characteristics. The adequate number of evaluators is one of the recurrent concerns related to the efficiency of its use. The objective of this research was to verify the implications of the number of evaluators on the visual evaluation efficiency for plant vigor in *U. ruziziensis* and determine the optimized number of evaluators. For this purpose, 2204 genotypes of *U. ruziziensis* were evaluated in the augmented block design using two controls: Marandu (*U. brizantha*) and Basilisk (*U. decumbens*). Green mass production (GMP) was measured and plant vigor grades were given by visual inspection carried out by six evaluators. Data were analyzed using mixed models approach with recovery of interblock and intergenotypic information. Regarding vigor grades, analyzes were performed considering each evaluator's grade and combination averages of 2, 3, 4, 5 and 6 evaluators. Variance components, heritability, selective accuracy and experimental variation coefficient (EVC) were estimated. The genetic variance was significant for both characteristics. It was verified that the indirect selection of productive genotypes through grades was efficient. The correlation between GMP and vigor grades for six evaluators was high (0.86). The number of evaluators employed had repercussions on selective accuracy, genetic variance and EVC estimates, being unfavorable when using a single evaluator. Based on the results, it is possible to state that the optimized number of evaluators for selecting *U. ruziziensis* genotypes for biomass production is three.

Key words: Plant selection, Visual evaluation, Number of evaluators, Accuracy, Correlation, Heritability, Forage breeding.

1. INTRODUÇÃO

A *Urochloa ruziziensis* (R. Germ. & CM Evrard) Crins. (sinonímia *Brachiaria ruziziensis*) é uma cultura forrageira de grande potencial agronômico e, especialmente, nutricional (VALLE et al. 2000; LOPES et al. 2010; SOUZA SOBRINHO et al. 2010). Além disso, por se tratar de uma espécie exclusivamente sexual e diploide, é bastante utilizada em programas de melhoramento (NDIKUMANA, 1985; SAVIDAN; JANK; VALLE, 1989; TIMBÓ et al. 2014, VALLE et al. 2009).

O processo seletivo em *U. ruziziensis* e outras espécies forrageiras tem adotado a inspeção visual para avaliar alguns caracteres de genótipos em campo, visto que a praticidade do emprego da nota para a seleção é constantemente reportada em estudos na literatura (BRUCKNER; RAYMER; BURTON, 1991; RIDAY, 2009). Dentre as características frequentemente avaliadas, pode-se destacar a resistência à pragas e doenças (SILVA, 2013), a densidade de perfilho (FIGUEIREDO; NUNES; VALLE, 2012), o florescimento, o hábito de crescimento (NGUYEN; SLEPER, 1983), a rebrota, a tolerância ao frio (FÈ, 2015), “stay green” (WILKINSON, 2003), o vigor da planta (TEIXEIRA, 2015) e a produção de massa verde (BURTON, 1974; BURTON, 1982; PRICE; CASLER, 2014).

A eficiência da seleção visual no melhoramento de espécies forrageiras depende de vários fatores dentre os quais a natureza do caráter alvo do melhoramento se destaca. Para caracteres quantitativos, a exemplo da produção de grãos, esta eficiência é dita baixa (ATKINS, 1964; ABADASSI, 2014; ABREU et. al, 2010; STUTHMAN; ATEIDL; 1976). Contudo, a avaliação visual para caracteres qualitativos nos programas de melhoramento da *U. ruziziensis* tem apresentado resultados satisfatórios (SOBRINHO, 2010).

Outro fator que pode afetar a eficiência da seleção visual é a quantidade de avaliadores empregados no campo. Informações sobre esse assunto em *U. ruziziensis* são escassos. Na literatura são retratados trabalhos com número variável de avaliadores (um, seis ou dez) (ABREU et. al, 2010; TEIXEIRA, 2015). Contudo, em nenhum destes foram apresentadas razões que justificassem essas quantidades.

A determinação do número otimizado de avaliadores pode contribuir para obtenção de estimativas paramétricas e predições de valores genéticos mais acurados, além de contribuir também para a viabilidade econômica no que concerne ao caso mais judicioso dos recursos humanos e financeiros. Essa determinação deve ser amparada por meio de parâmetros descritivos da correspondência da seleção relativo ao caráter de interesse (e.g. correlação) (PRICE; CASTLER, 2014), bem como medidas da qualidade da informação obtida a partir

dos dados experimentais (e.g. acurácia seletiva e coeficiente de variação) (BURTON, 1974; HENDERSON, 1984).

Portanto, o objetivo do presente trabalho foi verificar as implicações do número de avaliadores na eficiência da avaliação visual para produção de massa verde de genótipos de *U. ruziziensis* e determinar o número otimizado de avaliadores.

2. MATERIAL E MÉTODOS

2.1. Descrição do experimento

2.1.1. Local

O experimento foi conduzido no campo experimental da Embrapa Gado de Leite sediado na cidade de Coronel Pacheco, MG. O município encontra-se localizado a 410 metros de altitude, a 21°33' latitude sul e a 43°06' latitude oeste. O solo da área experimental é classificado como Argissolo Vermelho-Amarelo Álico (SANTOS et al., 2006). Segundo Köppen, o clima é classificado como tropical de altitude (tipo Cwa mesotérmico) com diminuição de chuvas no inverno e temperatura média anual de 19° C. Apresenta invernos secos e frios, ao passo que os verões são chuvosos com temperaturas moderadamente altas. A precipitação média anual é de 1536 mm.

2.1.2. Genótipos avaliados, delineamento e condução do experimento

Em agosto de 2011 foi instalado um experimento em blocos aumentados constituído de 51 blocos, os quais variaram entre 28 a 72 genótipos dentro de cada bloco. Os tratamentos regulares constaram de 2204 genótipos de *U. ruziziensis*, os quais foram obtidos por sementes provenientes do segundo ciclo de seleção recorrente intraespecífica do programa de melhoramento genético da Embrapa Gado de Leite. Os tratamentos comuns foram compostos por duas testemunhas: cultivar Marandu (*U. brizantha*) e cultivar Basilisk (*U. decumbens*). A parcela foi constituída de uma planta com espaçamento de um metro entre as parcelas.

No momento do plantio foi realizada uma adubação de 350kg/ha da fórmula 8-28-16 (NPK). Ao longo do ano foi realizado uma adubação em cobertura de 1ton/ha da fórmula 20-05-20, a qual foi particionada ao longo dos cortes no período chuvoso. Foram realizadas capinas manuais quando necessário.

No início de dezembro de 2011 foi realizado um corte de uniformização e em 2012 foram realizados dois cortes de avaliação, sendo um em janeiro e outro em fevereiro com intervalo de 40 dias aproximadamente. Para a execução do presente trabalho foi considerado apenas o primeiro corte de avaliação.

2.2. Características avaliadas

Foram avaliadas as características vigor associada à produtividade e produção de massa verde (PMV). Para avaliação do vigor foram utilizados seis avaliadores que, mediante inspeção visual, atribuíram notas aos genótipos de acordo com a seguinte escala: 1 – muito ruim, 2 – ruim, 3 – regular, 4 – bom e 5 – muito bom. Vale ressaltar que as notas foram atribuídas antes do corte ser realizado. Para avaliação da PMV, procedeu-se ao corte das plantas a 5,0 cm do nível do solo sendo a aferição realizada por meio de balança de suspensão portátil.

2.3. Análise estatística dos dados

Os dados da PMV e das notas foram analisados via abordagem de modelos mistos (HENDERSON, 1984) com recuperação de informação interblocos e intergenotípica de acordo com o modelo a seguir:

$$y = 1\mu + Z_1b + Z_2g + e \quad (1)$$

em que: y : vetor de dados da PMV; μ : intercepto; Z_1 : matriz do modelo referente aos efeitos aleatórios dos blocos; Z_2 : matriz do modelo referente aos efeitos aleatórios dos genótipos; b : vetor dos efeitos aleatórios dos blocos, sendo $b \sim NMV(0, I\sigma_b^2)$; g : vetor dos efeitos aleatórios dos genótipos, sendo $g \sim NMV(0, I\sigma_g^2)$ e e : vetor dos erros, sendo $e \sim NMV(0, I\sigma_e^2)$.

No caso da análise da característica vigor associada à produtividade, y correspondeu às notas de cada avaliador e às médias das combinações de 2, 3, 4, 5 e 6 avaliadores.

Para ambas as características, a estimação dos efeitos fixos (melhor estimador linear não tendencioso – BLUE) e a predição dos efeitos aleatórios (melhor predição linear não viciada – BLUP) foram realizadas via resolução do sistema de equações de Henderson (HENDERSON, 1975). As componentes da variância foram estimados pelo método da máxima verossimilhança residual (REML) (PATTERSON; THOMPSON, 1971). A

significância das componentes da variância foi avaliada pelo teste da razão de verossimilhança a 5%. Para isso, foi utilizado a função *lmer* do pacote *lme4* (BATES et al., 2015).

A partir das estimativas das componentes da variância foi estimada a herdabilidade em nível de parcela por meio da expressão: $h^2 = \hat{\sigma}_g^2 / \hat{\sigma}_f^2$, em que $\hat{\sigma}_g^2$: estimativa da variância genética e $\hat{\sigma}_f^2$: estimativa da variância fenotípica. Para a estimação da acurácia seletiva foi empregado a seguinte equação (HENDERSON, 1984): $\hat{r}_{gg} = (1 - \overline{PEV} / \sigma_g^2)^{1/2}$, em que \overline{PEV} : estimativa da variância do erro de predição médio. Ademais, também foram estimados o coeficiente de variação experimental e as correlações de Spearman entre os BLUPs dos genótipos para PMV e notas.

3. RESULTADOS

Observou-se variância genética significativa a um nível nominal de 5% para as características PMV e notas (TABELA 1). A herdabilidade em nível de parcela para a característica PMV foi de 0,35, com correspondente acurácia seletiva de 68%. Além disso, foi possível observar magnitudes elevadas de CVe (32%). A herdabilidade em nível de parcela para a nota média atribuída por seis avaliadores foi de 0,39 e acurácia seletiva de 0,62.

Tabela 1 – Estimativas da variância genética ($\hat{\sigma}_g^2$), variância ambiental ($\hat{\sigma}_e^2$), herdabilidade (\hat{h}^2), acurácia seletiva (\hat{r}_{gg}) e coeficiente de variação experimental (CVe) para as característica PMV e notas de vigor, na média de 6 avaliadores.

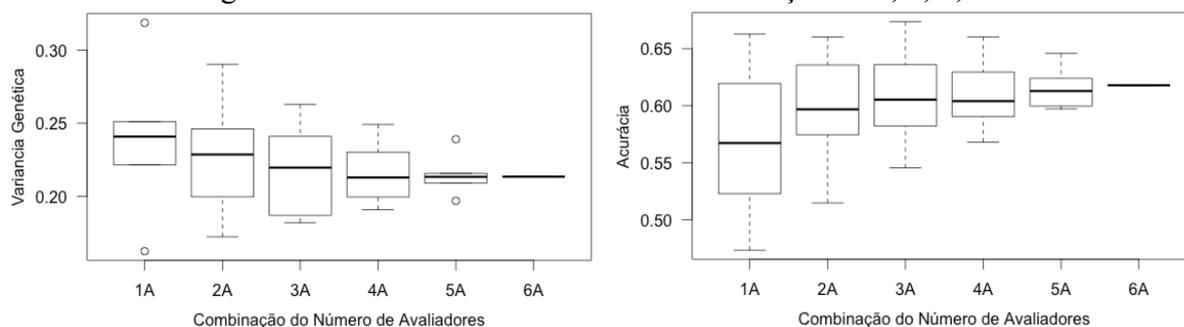
Parâmetro	PMV	Nota
		6 Avaliadores
σ_g^2	141849*	0.21*
σ_e^2	157948	0.34
h^2	0.35	0.39
r_{gg}	0.68	0.62
CVe	0.32	0.19

*Significativo pelo teste de razão de verossimilhança a 0,05%.

Notou-se grande flutuação nas estimativas dos parâmetros variância genética, acurácia seletiva e CVe (FIGURAS 2 e 3) nas diferentes combinações de números de avaliadores. A variação foi destacadamente maior quanto menos avaliadores foram empregados. Observou-

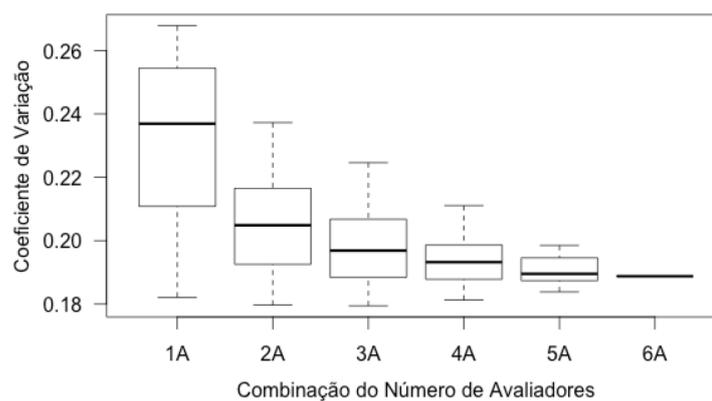
se que a média das estimativas se mantiveram mais estáveis a medida que o incremento do número de avaliadores foi maior que três.

Figura 1 – Boxplot das estimativas da variância genética e da acurácia seletiva para as notas de vigor de um avaliador e na média da combinação de 2, 3, 4, 5 e 6.



Fonte: Próprio autor.

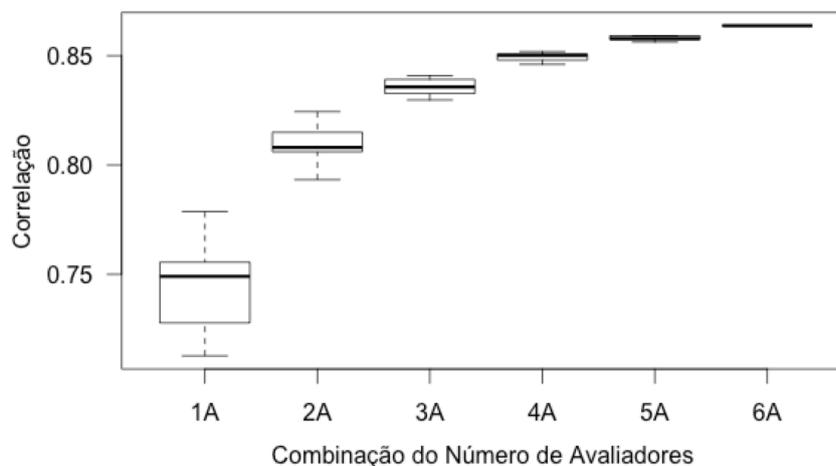
Figura 2 – Boxplot das estimativas do coeficiente de variação experimental (CVE) para notas de vigor de um avaliador e na média da combinação de 2, 3, 4, 5 e 6 avaliadores.



Fonte: Próprio autor.

A eficiência da nota para fins seletivos foi investigado nas diferentes combinações de número de avaliadores mediante a estimação das correlações entre os BLUPs para PMV e nota. Pela figura 3, constatou-se um incremento significativo nesta correlação à medida que o número de avaliadores aumentou, demonstrando a viabilidade do emprego da nota para selecionar indiretamente genótipos mais produtivos.

Figura 3 – Boxplot das correlações de Spearman entre os BLUPs dos genótipos para as características PMV e notas de vigor considerando um avaliador e media de 2, 3, 4, 5 e 6 avaliadores.



Fonte: Próprio autor.

4. DISCUSSÃO

A avaliação de caracteres por meio da inspeção visual tem sido praticada nos programas de melhoramento de espécies forrageiras nas diferentes etapas dos ciclos seletivos (BRUCKNER; RAYMER; BURTON, 1991; RIDAY, 2009 SOUZA SOBRINHO, 2010, FIGUEIREDO; NUNES; VALLE, 2012). Neste trabalho foi retratada a avaliação de genótipos de *U. ruziziensis* na etapa inicial de um dos ciclos do programa de melhoramento. Esta etapa se caracteriza pelo grande número de genótipos por ensaio, sendo o emprego de notas efetivo devido sua praticidade na fenotipagem e economia de recursos proporcionada.

Outro aspecto relacionado à avaliação de genótipos em etapa inicial de ciclos seletivos é o emprego de delineamentos em blocos aumentados (DBA). A utilização deste delineamento é indicado nestas fases devido a sua flexibilidade e, principalmente, por permitir testar genótipos não repetidos. Contudo, as estimativas das componentes de variância são, em geral, menos acuradas que as estimativas provenientes de outros delineamentos, a exemplo dos blocos completos (SANTOS et al., 2002). Figueiredo et al. (2013) apresentaram valores mais elevados de acurácia ao analisar a nota de rebrota em *Urochloa humidicola* num delineamento em blocos completos.

O fato das estimativas provenientes do DBA serem menos acuradas pode justificar, em partes, as baixas magnitudes das estimativas das variância genética para PMV e notas (TABELA 1). Mesmo assim, vale ressaltar que nesta etapa inicial, as diferenças a serem detectadas são mais elevadas, de modo que o uso do DBA é adequado.

Vários fatores podem afetar a eficiência da avaliação visual. Os resultados obtidos mostraram que o número de avaliadores tem implicações sobre os parâmetros variância genética, acurácia seletiva e CVe, sendo esta mais marcante para o caso de um único avaliador. Ademais, foi constatado que os casos com três ou mais avaliadores constituíram um cenário mais adequado para se proceder a avaliação por meio de notas.

A escala de notas atribuída para aferir uma determinada característica deve ser avaliada, ou demonstrada sua correspondência para fins seletivos. Foi possível identificar forte correspondência na classificação dos genótipos pelos BLUPs da PMV e notas e que esta correlação teve incremento expressivo com o aumento de avaliadores. Souza Sobrinho et al. (2010) destacou a eficiência da avaliação visual para a característica resistência à cigarrinhas demonstrando a utilidade da nota para avaliar este caráter. Portanto, a avaliação visual se constitui numa estratégia interessante para fins de melhoramento genético de *U. ruziziensis* para avaliar a produção de massa verde de forma indireta.

5. CONCLUSÃO

O número de avaliadores tem implicações na eficiência da avaliação visual para produção de massa verde em *U. ruziziensis*. Para fins de otimização, recomenda-se o emprego de três avaliadores.

REFERÊNCIAS

- ABADASSI, J., Mass Selection to Improve Grain Yield in a Traditional Tropical Maize Population. **International Journal of Science and Technology**, v. 4, n. 3, 2014.
- ABREU, G. B.; RAMALHO, M. A. P.; TOLEDO, F. H. R. B.; SOUZA, J. C. de. Strategies to improve mass selection in maize. **Maydica**, Bergamo, v. 55, n. 4, p. 219-225, 2010.
- ATKINS, R. E. Visual Selection for Grain Yield in Barley. **Crop Science**, v. 4, n. 5, p. 494-497, 1964.
- BRUCKNER, P. L.; RAYMER, P. L.; BURTON, G. W. Recurrent phenotypic selection for forage yield in rye. **Euphytica**, v. 54, n. 1, p. 11-17, 1991.
- BURTON, G. W. Improved recurrent restricted phenotypic selection increases bahiagrass forage yields. **Crop Science**, v. 22, n. 5, p. 1058-1061, 1982.
- BURTON, G. W. Recurrent restricted phenotypic selection increases forage yields of *Pensacola bahiagrass*. **Crop Science**, v. 14, n. 6, p. 831-835, 1974.
- FÈ, D. et al. Genetic and environmental variation in a commercial breeding program of perennial ryegrass. **Crop Science**, v. 55, n. 2, p. 631-640, 2015.
- FIGUEIREDO, U. J.; NUNES, J. A. Rodrigues; VALLE, C. B. do. Estimation of genetic parameters and selection of *Brachiaria humidicola* progenies using a selection index. **Crop Breeding and Applied Biotechnology**, v. 12, n. 4, p. 237-244, 2012.
- HENDERSON, C. R. **Applications of linear models in animal breeding**. 384 p. 1984.
- LOPES, F. C. et al. Composição química e digestibilidade ruminal in situ da forragem de quatro espécies do gênero *Brachiaria*. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 62, n. 4, p. 883-888, 2010.
- NDIKUMANA, J. Etude de l'hybridation entre espèces apomictiques et sexuées dans le genre *Brachiaria*. **Ph.D Thesis**. Université Catholique de Louvain, Louvain-la-Neuve, Belgique. 210 p. 1985.
- NGUYEN, H. T.; SLEPER, D. A. Genetic variability of seed yield and reproductive characters in tall fescue. **Crop science**, v. 23, n. 4, p. 621-626, 1983.
- PRICE, D. L.; CASLER, M. D. Divergent selection for secondary traits in upland tetraploid switchgrass and effects on sward biomass yield. **BioEnergy Research**. 2014.
- RIDAY, H. Correlations between visual biomass scores and forage yield in space planted red clover (*Trifolium pratense L.*) breeding nurseries. **Euphytica**, v. 170, n. 3, p. 339-345, 2009.
- Santos, A. H., B., E., Ferreira, D. F., Silva Filho, J. L. D. Simulation of mixed models in augmented block design. **Scientia Agricola**, 59(3), 483-489. 2002.
- SAVIDAN, Y. H., JANK, L., COSTA, J. C. G.; VALLE, C. D. Breeding *Panicum maximum* in Brazil. 1. Genetic resources, modes of reproduction and breeding procedures. **Euphytica**, v. 41, n. 1, p. 107-112, 1989.

SOUZA SOBRINHO, F. de et al. Repetibilidade de características agronômicas e número de cortes necessários para seleção de *Urochloa ruziziensis*. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 6, p. 579-584, jun. 2010.

STUTHMAN, D. D.; STEIDL, R. P. Observed gain from visual selection for yield in diverse oat populations. **Crop Science**, v. 16, n. 2, p. 262-264, 1976.

TIMBO, A. L. de O. et al. Obtaining tetraploid plants of ruzigrass (*Brachiaria ruziziensis*). **R. Bras. Zootec.**, Viçosa, v. 43, n. 3, p. 127-131, Mar. 2014.

VALLE, C.B. do; EUCLIDES, V. P. B.; MACEDO, M. C. M. Características das plantas forrageiras do gênero *Brachiaria*. SIMPÓSIO SOBRE MANEJO DA PASTAGEM, v. 17, p. 65-108, 2000.

WILKINSON, J. M.; HILL, J. Effect on yield and dry-matter distribution of the stay-green characteristic in cultivars of forage maize grown in England. **Grass and Forage Science**, v. 58, n. 3, p. 258-264, 2003.

ARTIGO 2

MODELOS DE ANÁLISE DE NOTAS NO MELHORAMENTO DE *Urochloa ruzizensis*

RESUMO

A *Urochloa ruzizensis* (sin. *Brachiaria ruzizensis*) é uma espécie forrageira importante por apresentar características nutritivas desejáveis e, principalmente, por ser a única espécie do gênero exclusivamente sexual e diploide. Em muitas situações, sua avaliação é efetuada por meio da atribuição de notas aos genótipos. Por essa razão, o presente trabalho teve por objetivo comparar quatro alternativas de análise de notas, a citar: abordagem de modelos lineares mistos na escala original das notas, modelos lineares mistos na escala da nota transformada via Box-Cox, modelo linear misto generalizado e modelo linear misto generalizado bayesiano. Ademais, analisar os dados da produção de massa verde com a finalidade de proceder ao estudo da correspondência dos genótipos selecionados por meio das notas usando as diferentes alternativas. Para isso, 2204 genótipos de *U. ruzizensis* foram avaliados em um experimento no delineamento em blocos aumentados. Seis avaliadores foram utilizados para atribuir notas de vigor aos genótipos. As alternativas foram apresentadas em cenários distintos, considerando a nota de um, de três e de seis avaliadores. Para os critérios de comparação foram utilizados a acurácia seletiva, a correlação de Spearman, a eficiência de seleção e o ganho com a seleção. A variância genética para ambas características foi significativa. Foi possível validar a seleção indireta para produção de massa verde dos genótipos em campo por meio da atribuição das notas de vigor. Não foi possível detectar diferenças significativas entre as alternativas de análise das notas, porém deve-se atentar para a adequação dos modelos lineares mistos generalizados para análise de notas na sua escala original.

Palavra-chave: Análise visual, Modelo linear misto generalizado, Transformação Box-Cox, Melhoramento genético, Inferência bayesiana

ABSTRACT

Urochloa ruziziensis (syn. *Brachiaria ruziziensis*) is an important forage species because it has desirable nutritional characteristics and, above all, because it is the only specie among *Urochloa ssp* to be exclusively sexual and diploid. In many situations, assigning grades to genotypes does its evaluation. For this reason, this study intends to compare four alternatives to analyze grades: linear mixed model in original grade scale, linear mixed model in transformed Box-Cox scale, generalized linear mixed model and bayesian generalized linear mixed model. In addition, to analyze genotype green mass production data in order to proceed a study about genotypes correspondence through grade evaluation using different alternatives of analysis, we assessed 2204 *U. ruziziensis* genotypes in an augmented block design experiment. Six evaluators were used to assign vigor grades to genotypes. Alternatives were presented in different scenarios, considering grades given by one, three and six evaluators. Selective accuracy, Spearman correlation, selection efficiency and selection gain were used for comparisons criteria. Genetic variance for both characteristics was significant. It was possible to validate the indirect selection of the genotypes in the field by assigning vigor grades. It was not possible to detect significant differences between grade alternative analyses, but attention should be paid to the adequacy of generalized linear mixed models in regards of their grade analysis on original scale.

Key words: Visual analysis, Generalized linear mixed model, Box-Cox Transformation, Genetic improvement, Bayesian inference

1. INTRODUÇÃO

A *Urochloa ruziziensis* (R. Germ. & CM Evrard) Crins. (sinonímia *Brachiaria ruziziensis*) tem despertado interesse entre os pesquisadores envolvidos no melhoramento genético de culturas forrageiras devido ao seu potencial agrônomo e, sobretudo, nutricional (VALLE et al. 2000; LOPES et al. 2010; SOUZA SOBRINHO et al. 2010), sua capacidade de participar de cruzamentos intra e interespecíficos (NDIKUMANA, 1985; SAVIDAN; JANK; VALLE, 1989; TIMBÓ et al. 2014) e sua importância quando se deseja avançar gerações em programas de melhoramento que optem por realizar o método de seleção recorrente (VALLE et al. 2009).

Estudos relacionados ao melhoramento de forrageiras têm reportado a eficiência da inspeção visual para proceder as avaliações em campo e promover a seleção dos genótipos superiores (BRUCKNER; RAYMER; BURTON, 1991; RIDAY, 2009). Dentre as características frequentemente avaliadas por meio da atribuição de notas, pode-se destacar a resistência à pragas e doenças (SILVA, 2013), a densidade de perfilho (FIGUEIREDO; NUNES; VALLE, 2012), o florescimento, o hábito de crescimento (NGUYEN; SLEPER, 1983), a rebrota, a tolerância ao frio (FÈ, 2015), “stay green” (WILKINSON, 2003), o vigor da planta (TEIXEIRA, 2015) e produção de massa verde (BURTON, 1974; BURTON, 1982; PRICE; CASLER, 2014),

A maneira de proceder as análises estatísticas de notas tem gerado questionamentos entre os profissionais do melhoramento vegetal. Embora a escala das notas seja classificada como variável discreta ordinal (AGRESTI, 2002), sua análise tem sido frequentemente realizada por meio de modelos que pressupõem distribuição normal, apoiando-se no teorema do limite central (MOOD; GRAYBILL; BOES, 1974) ou por meio da transformação na escala da nota (BOX; COX, 1964).

A adoção de modelos inadequados para análise de dados, a exemplo das notas, pode resultar em estimativas que não representem a verdade sobre os parâmetros, produzindo resultados espúrios (JAEGER, 2008; BOLKER, 2009; STROUP, 2015). No contexto do melhoramento genético de plantas, tal inadequação pode culminar em um ordenamento errado dos genótipos superiores (erro tipo 3), permitindo a eliminação de genótipos promissores, bem como a seleção de genótipos de menor performance.

Existem alternativas mais apropriadas para analisar notas, a exemplo dos modelos lineares mistos generalizados (MLMG). Estes modelos são uma extensão do modelo linear misto (MLM) por englobar distribuições da família exponencial, da qual a multinomial é

frequentemente utilizada para descrever variáveis categóricas ordinais (e.g. notas). Modelos multinomiais podem ser empregados sob enfoque frequentista (BRESLOW, CLAYTON, 1993; MCCULLOCH, 1997) ou sob enfoque bayesiano (GIANOLA & FOULLEY, 1983; GIANOLA & FERNANDO, 1986). Quando se utiliza os modelos lineares mistos generalizados bayesianos (MLMGB) as estimativas paramétricas são revisadas pelo teorema de Bayes (BAYES; PRICE, 1763), o qual afirma que a probabilidade *a posteriori* da ocorrência de um evento é proporcional a verossimilhança da ocorrência desse evento ponderado por um conhecimento *a priori*. Gouy et al. (2013) ressaltaram a adequação do MLMGB (threshold) para analisar as notas da reação de resistência à ferrugem na seleção de clones de cana-de-açúcar.

Diante do exposto, no presente trabalho objetivou-se comparar alternativas de análise de notas de vigor provenientes da avaliação visual de genótipos de *U. ruziziensis* pela abordagem de MLM na escala original das notas, MLM na escala da nota transformada via Box-Cox, MLMG e MLMGB. Além disso, analisar os dados da produção de massa verde dos genótipos utilizando MLM, com a finalidade de proceder ao estudo da correspondência dos genótipos selecionados por meio das notas.

2. MATERIAL E MÉTODOS

2.1. Descrição do experimento

2.1.1. Local

O experimento foi conduzido no campo experimental da Embrapa Gado de Leite sediada na cidade de Coronel Pacheco, MG. O município encontra-se localizado a 410 metros de altitude, a 21°33' latitude sul e a 43°06' latitude oeste. O solo da área experimental é classificado como Argissolo Vermelho-Amarelo Álico (SANTOS et al., 2006). Segundo Köppen, o clima é classificado como tropical de altitude (tipo Cwa mesotérmico) com diminuição de chuvas no inverno e temperatura média anual de 19° C. Apresenta invernos secos e frios, ao passo que os verões são chuvosos com temperaturas moderadamente altas. A precipitação média anual é de 1536 mm.

2.1.2. Genótipos avaliados, delineamento e condução do experimento

Em agosto de 2011 foi instalado um experimento em blocos aumentados constituído de 51 blocos, os quais variaram entre 28 a 72 genótipos dentro de cada bloco. Os tratamentos regulares constaram de 2204 genótipos de *U. ruziziensis*, os quais foram obtidos por sementes provenientes do segundo ciclo de seleção recorrente intraespecífica do programa de melhoramento genético da Embrapa Gado de Leite. Os tratamentos comuns foram compostos por duas testemunhas: cultivar Marandu (*U. brizantha*) e cultivar Basilisk (*U. decumbens*). A parcela foi constituída de uma planta com espaçamento de um metro entre as parcelas.

No momento do plantio foi realizada uma adubação de 350kg/ha da fórmula 8-28-16 (NPK). Ao longo do ano foi realizado uma adubação em cobertura de 1ton/ha da fórmula 20-05-20, a qual foi particionada ao longo dos cortes no período chuvoso. Foram realizadas capinas manuais quando necessário.

No início de dezembro de 2011 foi realizado um corte de uniformização e em 2012 foram realizados dois cortes de avaliação, sendo um em janeiro e outro em fevereiro com intervalo de 40 dias aproximadamente. O presente trabalho considerou apenas o primeiro corte de avaliação para simplificar a apresentação das alternativas de análise de notas.

2.2. Características avaliadas

Foram avaliadas as características vigor associada à produtividade e produção de massa verde (PMV). Para avaliação do vigor foram utilizados seis avaliadores que, mediante inspeção visual, atribuíram notas aos genótipos de acordo com a seguinte escala: 1 – muito ruim, 2 – ruim, 3 – regular, 4 – bom e 5 – muito bom. Vale ressaltar que as notas foram atribuídas antes do corte ser realizado. Para avaliação da PMV, procedeu-se ao corte das plantas a 5,0 cm do nível do solo sendo a aferição, em gramas, realizada por meio de balança de suspensão portátil.

2.3. Análise estatística dos dados

2.3.1. Análise dos dados da PMV

Os dados da PMV foram analisados de acordo com o seguinte MLM:

$$y = 1\mu + Z_1b + Z_2g + e \quad (1)$$

em que: y : vetor de dados da PMV; μ : intercepto; Z_1 : matriz do modelo referente aos efeitos aleatórios dos blocos; Z_2 : matriz do modelo referente aos efeitos aleatórios dos genótipos; b :

vetor dos efeitos aleatórios dos blocos, sendo $b \sim NMV(0, I\sigma_b^2)$; g : vetor dos efeitos aleatórios dos genótipos, sendo $g \sim NMV(0, I\sigma_g^2)$; e : vetor dos erros, sendo $e \sim NMV(0, I\sigma_e^2)$.

A estimação dos efeitos fixos (melhor estimador linear não tendencioso – BLUE) e a predição dos efeitos aleatórios (melhor predição linear não viciada – BLUP) foram realizadas via resolução do sistema de equações de Henderson (HENDERSON, 1975). As componentes da variância foram estimadas pelo método da máxima verossimilhança residual (REML) (PATTERSON; THOMPSON, 1971). A significância das componentes da variância foi avaliada pelo teste da razão de verossimilhança. A deviance do modelo foi estimada segundo a equação $D = -2\log l$, em que l corresponde a função de verossimilhança do modelo. A análise foi realizada com auxílio da função *lmer* do pacote *lme4* (BATES et al., 2015) (APÊNDICE A).

Para a estimação da herdabilidade em nível de indivíduo foi empregado a seguinte equação:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_b^2 + \sigma_e^2} \quad (2)$$

em que: σ_g^2 : variância genética; σ_b^2 : variância de bloco; σ_e^2 : variância residual.

2.3.2. Alternativas de análise das notas de vigor associada à produtividade

As notas foram analisadas por quatro alternativas diferentes, a citar: abordagem de MLM na escala original das notas, MLM na escala da nota transformada via Box-Cox, MLMG e MLMGB, em três cenários descritos a seguir.

2.3.2.1. Cenário 1 – notas de um avaliador

Foram consideradas apenas as notas atribuída por um avaliador tomado aleatoriamente. Para análise das notas na abordagem de MLM na escala original das notas foram utilizados os procedimentos descritos em 2.3.1. No caso da alternativa que aborda o MLM na nota transformada foi empregado a transformação Box-Cox, conforme a seguinte expressão:

$$y_t(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0 \\ \log y, & \text{se } \lambda = 0 \end{cases} \quad (3)$$

em que: λ : parâmetro que define a transformação; y : nota; $y_t(\lambda)$: nota transformada.

A estimação do valor de lambda (λ) baseou-se no método da maximização da função de verossimilhança (BOX; COX, 1964) do modelo (1) ajustado para a variável nota, tratando os efeitos do modelo como fixos. Para isso, foi utilizado a função *boxcox* do pacote *MASS* (VENABLES; RIPLEY, 2002). As notas transformadas foram ajustadas pelo modelo (1), seguindo os mesmos procedimentos descritos em 2.3.1.

A normalidade dos resíduos foi testada antes e após a transformação das notas por meio da análise gráfica do Q-Q plot da normal associado ao teste de Shapiro-Wilk (SHAPIRO; WILK, 1965). Para isso, foi utilizado as respectivas funções *qqnorm* e *Shapiro.test*, ambas do pacote *stats* (R CORE TEAM, 2015) (APÊNDICE A).

Para a abordagem do MLMG sob enfoque frequentista, as notas foram ajustadas por um modelo threshold, admitindo distribuição multinomial (AGRESTI, 2002; AGRESTI, 2007) e função de ligação probit, conforme descrito a seguir:

$$\Phi^{-1}(\pi_c | b, g) = 1\mu + Z_1b + Z_2g \quad (4)$$

em que: Φ^{-1} : função de ligação probit; π_c : probabilidade das categorias, sendo $c = 5$.

As estimativas de máxima verossimilhança dos parâmetros e as predições BLUPs dos efeitos genotípicos foram obtidas via aproximação Laplace da função de verossimilhança (PINHEIRO; CHAOS, 2006). Para isso, foi utilizado a função *clmm* do pacote *ordinal* (CHRISTENSEN, 2010). A significância das componentes da variância foi avaliada pelo teste da razão de verossimilhança. A deviance do modelo foi estimada. Para a estimação da herdabilidade, utilizou-se a expressão (2) sendo a σ_e^2 substituída pela σ_Φ^2 (APÊNDICE A).

Para a alternativa que aborda o MLMGB, as notas foram ajustadas segundo o modelo (4) com inclusão de priori gama inversa para as variâncias de bloco e genótipo, priori não informativa para os efeitos fixos e priori fixa de valor 1 para a variância residual (HADFIELD, 2012). Foram simulados 500000 iterações pelo método Monte Carlo via Cadeias de Markov (MCMC) com um *burn-in* de 100000 amostras e um *thin* a cada 200 iterações, resultando numa distribuição *a posteriori* de 2000 amostras de parâmetros estimados. A partir dessas amostras foram determinadas as médias e respectivos intervalos de credibilidade (HPD) para cada parâmetro. Para isso, foi utilizado a função *MCMCglmm* do

pacote *MCMCglmm* (HADFIELD, 2010). A herdabilidade foi estimada da mesma maneira que no MLMG (APÊNDICE A).

2.3.2.2. Cenário 2 – notas de três avaliadores

Foram consideradas as notas atribuídas por três avaliadores tomados ao acaso. Para as alternativas dos MLM nas escalas original e transformada das notas, as análises foram efetuadas pelo ajuste do modelo (1), sendo o vetor y correspondente às médias das notas dos três avaliadores. No caso da análise pelas alternativas dos MLMG e MLMGB, foi empregado o modelo (4) com o acréscimo do efeito fixo de avaliador e efeito aleatório de parcela, sendo $p \sim NMV(0, I\sigma_p^2)$. Os procedimentos de estimação-predição foram os descritos em 2.3.2.1. A estimativa da herdabilidade a nível de indivíduo, para as alternativas MLMG e MLMGB, foi obtida a partir do seguinte estimador:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_b^2 + \sigma_p^2 + \sigma_\Phi^2} \quad (5)$$

Em que: σ_p^2 : variância de parcela; σ_Φ^2 : variância da distribuição probit, cujo valor é 1.

2.3.2.3. Cenário 3 – notas de seis avaliadores

Para as alternativas dos MLM, foi adotado o modelo (1) acrescido do efeito fixo de avaliador e do efeito aleatório de parcela, sendo $p \sim NMV(0, I\sigma_p^2)$. Os procedimentos de análise foram descritos em 2.3.2.1. Para a estimação da herdabilidade a nível de indivíduo foi empregado a seguinte expressão:

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_b^2 + \sigma_p^2 + \sigma_e^2} \quad (6)$$

Para as alternativas dos MLMG e MLMGB, os procedimentos adotados foram descritos em 2.3.2.2.

2.4. Critérios para comparar as alternativas de análise das notas nos três cenários e suas correspondências com a análise da PMV

2.4.1. Acurácia seletiva

Nas alternativas dos MLM nas escalas original e transformada e MLMG procedeu-se a estimação da acurácia seletiva pela seguinte expressão (RESENDE, 2002):

$$\hat{r}_{\hat{g}g} = \left(1 - \frac{\overline{PEV}}{\hat{\sigma}_g^2} \right)^{1/2} \quad (7)$$

em que: \overline{PEV} : variância do erro de predição médio dos valores genotípicos dos tratamentos regulares; $\hat{\sigma}_g^2$: estimativa da variância genética.

Para a alternativa que emprega o MLMGB, a acurácia seletiva foi estimada pela raiz quadrada da herdabilidade (RESENDE, 2007).

2.4.2. Correlação e eficiência de seleção

Foram estimadas as correlações de Spearman entre os BLUPs dos genótipos para as diferentes alternativas de análise das notas e também para PMV. Para isso, empregou-se a função *cor* do pacote *stats* (R CORE TEAM, 2015) (APÊNDICE A).

Para estimação da eficiência de seleção (ES) entre as alternativas de análise das notas e PMV foi aplicado a expressão proposta por Hamblin e Zimmermann (HAMBLIN; ZIMMERMANN, 1986):

$$ES = \frac{A - C}{B - C} \times 100 \quad (8)$$

em que: A: número de genótipos selecionados e coincidentes, com base nos BLUPs, entre duas alternativas de análise de notas ou entre cada alternativa e a PMV; B: número de genótipos selecionados, admitindo uma proporção de seleção de 5%; C: número de genótipos coincidentes devido ao acaso ($C = B \times 0,05$).

2.4.3. Ganho com a seleção

Foram estimados os ganhos de seleção diretos e indiretos para a PMV com base na média dos BLUPs dos 5% melhores genótipos. Os ganhos indiretos foram determinados com base nos genótipos selecionados pelas diferentes alternativas de análises das notas. Além disso, foram estimados os ganhos com a seleção dos genótipos estatisticamente superiores, isto é, que apresentaram intervalos de predição acima da média geral (APÊNDICE A).

3. RESULTADOS

A variância genética foi significativa para as características PMV e notas de vigor para todas as alternativas de análise e cenários estudados (TABELA 1). A herdabilidade em nível de indivíduo para a característica PMV foi de 0,35, com correspondente acurácia seletiva de 68%. As magnitudes da herdabilidade para a nota de vigor flutuaram entre as alternativas de análise, a exemplo de 0,32 para notas norm e de 0,78 para notas BC, em contrapartida, essa flutuação não foi tão marcante entre os cenários, com exceção das notas BC no cenário 3 de seis avaliadores. Além disso, a acurácia seletiva apresentou valores médio a alto, tendo registrado 0,47 para notas MCMC e 0,92 para notas BC.

Com relação aos threshold, observou-se valores similares para os cenário 2 e 3, ao passo que, para o cenário 1, esses valores foram menores. A magnitude da variável latente mostrou-se quase inalterada entre os cenários 1 e 2. Pela deviance dos modelos, notou-se que o MLMGB foi o que mais se ajustou aos dados, enquanto o MLM sob nota transformada foi o de menor ajuste (TABELA 1).

Pela figura 1, evidenciou-se violação do pressuposto de normalidade dos resíduos do MLM na escala original da nota de vigor em todos os cenários. Ademais, verificou-se que a aplicação da transformação Box-Cox somente assegurou normalidade dos erros no cenário 3, cujo teste de Shapiro-Wilk acusou um p-valor superior ao nível nominal de significância adotado (5%).

Tabela 1 – Estimativas da variância genética ($\hat{\sigma}_g^2$), variância ambiental ($\hat{\sigma}_e^2$), herdabilidade em nível de indivíduo (\hat{h}^2), acurácia seletiva (\hat{r}_{gg}), variável latente, threshold e deviance para as características PMV e notas de vigor, retratada nos 3 cenários adotados (continua).

CENÁRIO 1 (um avaliador)							
Parâmetro	PMV	Notas norm	Notas BC	Notas CLMM	Notas MCMC		
					LI 2,5%	média a posteriori	LS 2,5%
σ_g^2	141849,20*	0,32*	180,92*	0,55*	0,25	0,56	1
σ_e^2	157947,60	0,50	28,03	1	1	1	1
h^2	0,35	0,32	0,78	0,29	0,16	0,28	0,41
r_{gg}	0,68	0,62	0,92	0,56	0,40	0,53	0,64
Variável latente	-	-	-	-	1,62	1,93	2,23
Threshold 1	-	-	-	-1,93	1,31	1,50	1,73

Continuação							
Threshold 2	-	-	-	-0,43	2,55	2,90	3,29
Threshold 3	-	-	-	0,97	3,71	4,23	4,76
Threshold 4	-	-	-	2,30	-	-	-
Deviance	35694	6145,9	18764	6082,6	-	5130,09	-
CENÁRIO 2 (três avaliadores)							
Parâmetro	Notas norm	Notas BC	Notas CLMM	Notas MCMC			
				LI 2,5%	média a posteriori	LS 2,5%	
σ_g^2	0,22*	3590,53*	1,31*	0,78	1,29	1,84	
σ_e^2	0,39	727,69	1	1	1	1	
h^2	0,31	0,74	0,27	0,16	0,26	0,37	
r_{gg}	0,59	0,90	0,58	0,40	0,51	0,61	
Variável latente	-	-	-	3,00	3,25	3,52	
Threshold 1	-	-	-3,27	2,33	2,46	2,59	
Threshold 2	-	-	-0,80	4,70	4,86	5,04	
Threshold 3	-	-	1,62	6,66	6,88	7,10	
Threshold 4	-	-	3,64	-	-	-	
Deviance	5471,2	25767	14371,4	-	9778,92	-	
CENÁRIO 3 (seis avaliadores)							
Parâmetro	Notas norm	Notas BC	Notas CLMM	Notas MCMC			
				LI 2,5%	média a posteriori	LS 2,5%	
σ_g^2	0,21*	0,12*	0,96*	0,54	0,94	1,37	
σ_e^2	0,28	0,18	1	1	1	1	
h^2	0,25	0,23	0,23	0,13	0,23	0,32	
r_{gg}	0,62	0,58	0,56	0,36	0,47	0,57	
Variável latente	-	-	-	2,99	3,21	3,40	
Threshold 1	-	-	-3,21	2,40	2,48	2,55	
Threshold 2	-	-	-0,73	4,57	4,67	4,76	
Threshold 3	-	-	1,46	6,39	6,52	6,65	
Threshold 4	-	-	3,31	-	-	-	
Deviance	27577	21294	26460	-	20680,09	-	

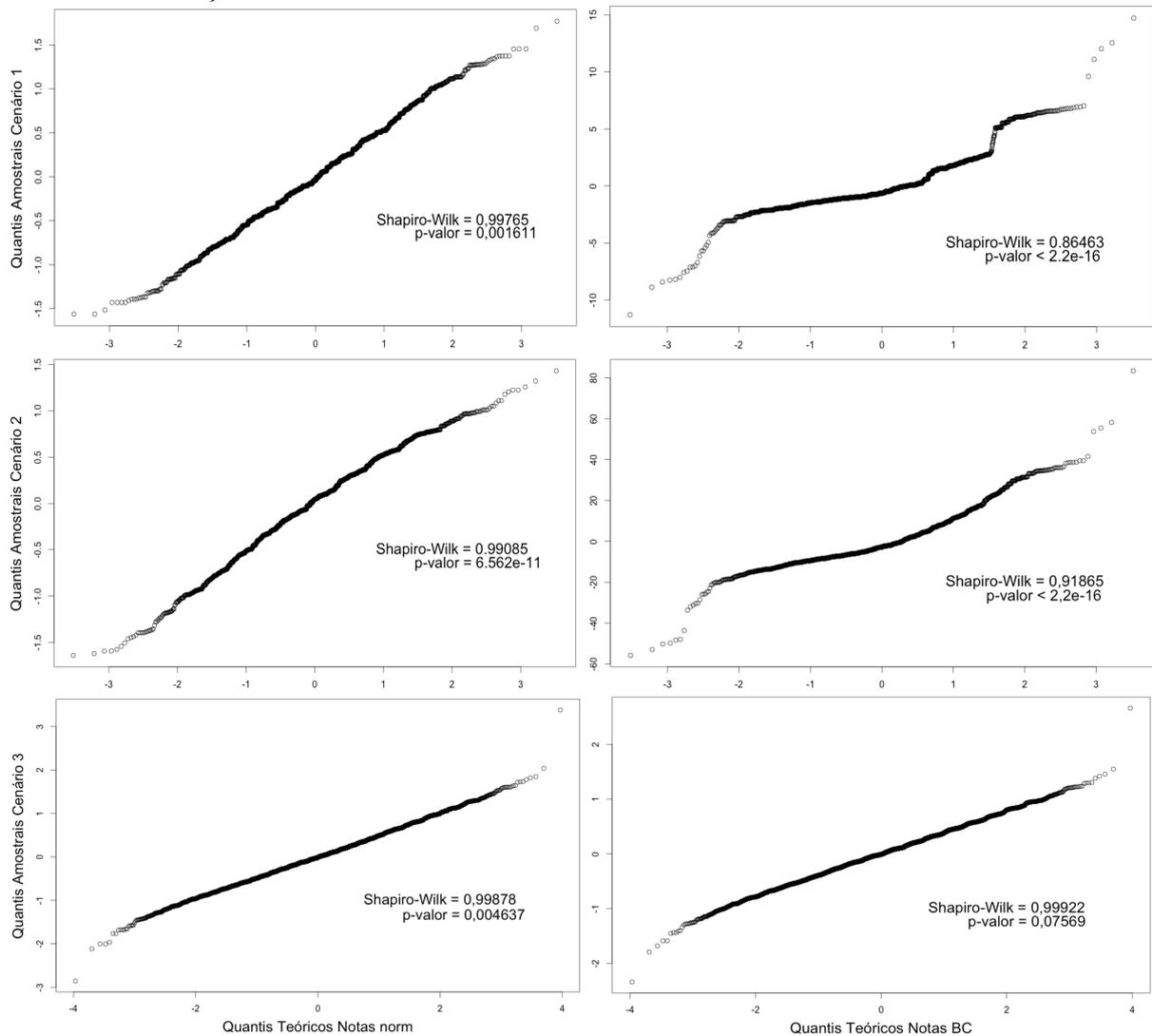
Notas norm: alternativa de análise sob suposição de normalidade na escala das notas; Notas BC: alternativa de análise considerando a nota na escala transformada via Box-Cox; Notas CLMM: alternativa de análise considerando a nota na escala original na abordagem de MLMG; Notas MCMC: alternativa de análise considerando a nota na escala original na abordagem de MLMGB.

LI: limite inferior.

LS: limite superior.

*Significativo pelo teste de razão de verossimilhança a 5%.

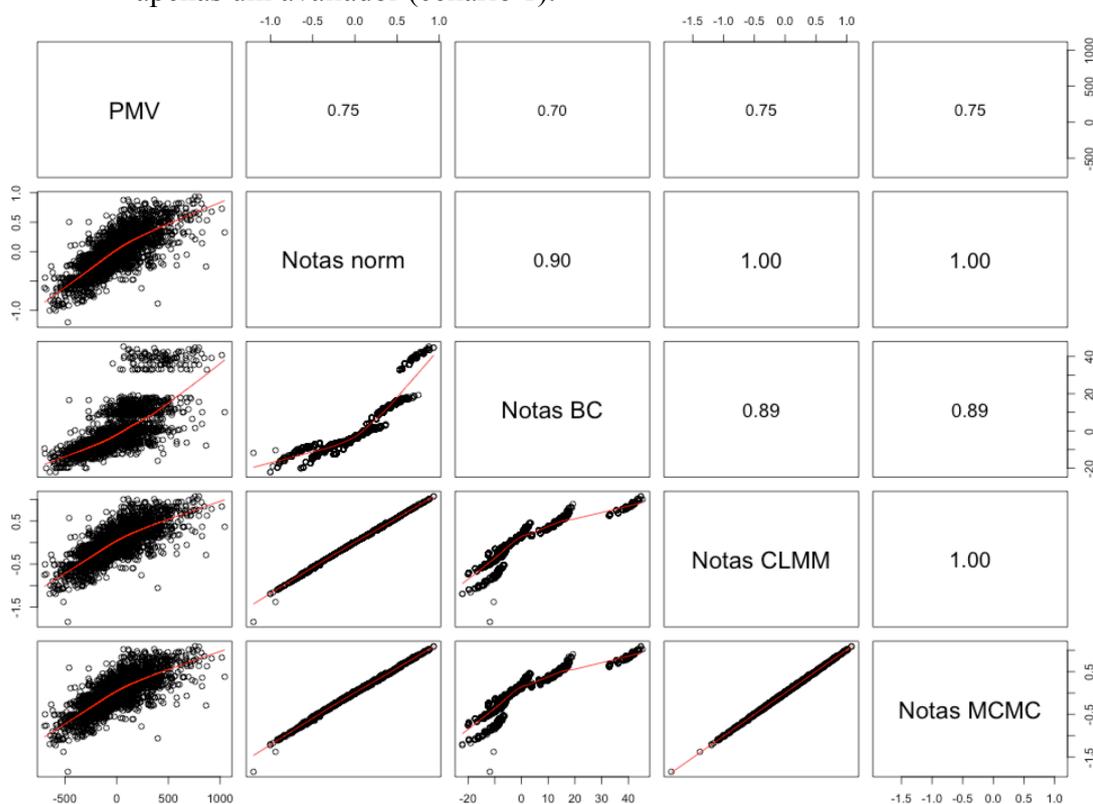
Figura 1: Q-Q plot da normal e teste de Shapiro-Wilk dos resíduos do MLM nas escalas original e transformada via Box-Cox da nota de vigor nos três cenários de avaliação.



Fonte: próprio autor.

As correlações dos BLUPs dos genótipos entre as alternativas variaram de 0,89 a 1 nos três cenários avaliados (FIGURAS 2, 3 e 4). Verificou-se correlações consistentemente mais elevadas com o acréscimo de avaliadores. Vale destacar as correlações de 1 entre as alternativas de ajuste dos modelos lineares mistos generalizados. Além disso, constatou-se um incremento nas estimativas das correlações entre os BLUPs para a característica PMV e notas com o aumento do numero de avaliadores. Contudo, esse aumento não se mostrou expressivo quando se considera três (cenário 2) ou mais avaliadores (cenário 3).

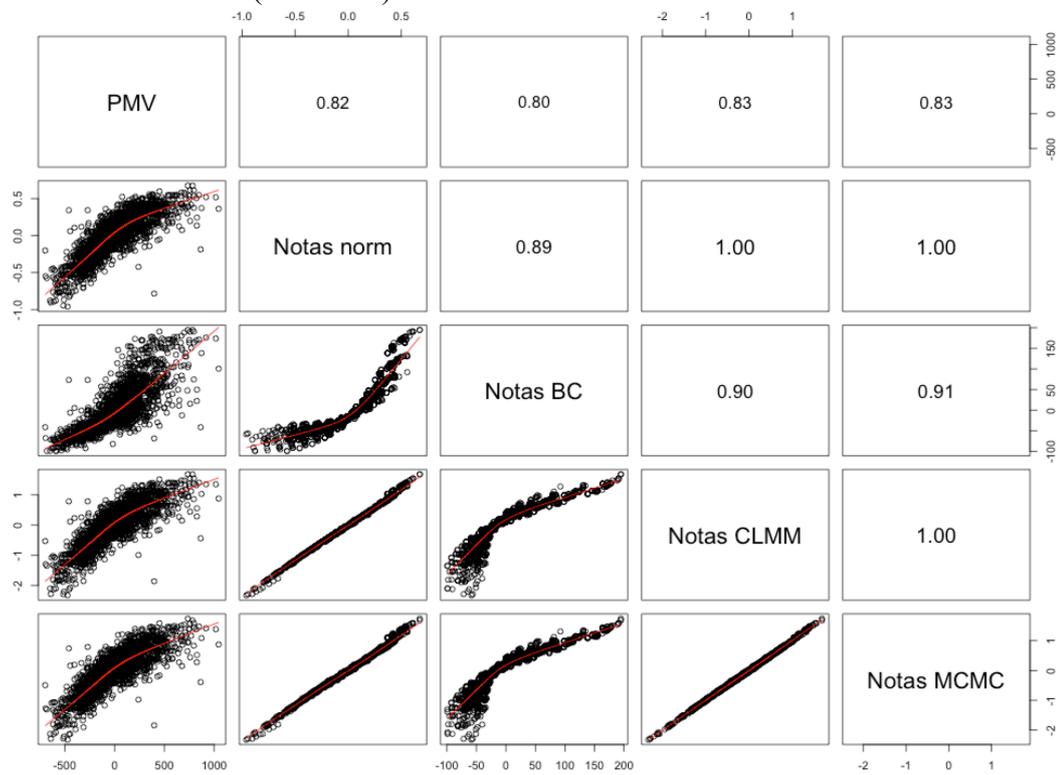
Figura 2 – Painel das correlações de Spearman entre BLUPs dos genótipos para a características PMV e nota de vigor nas diferentes alternativas considerando apenas um avaliador (cenário 1).



Fonte: próprio autor.

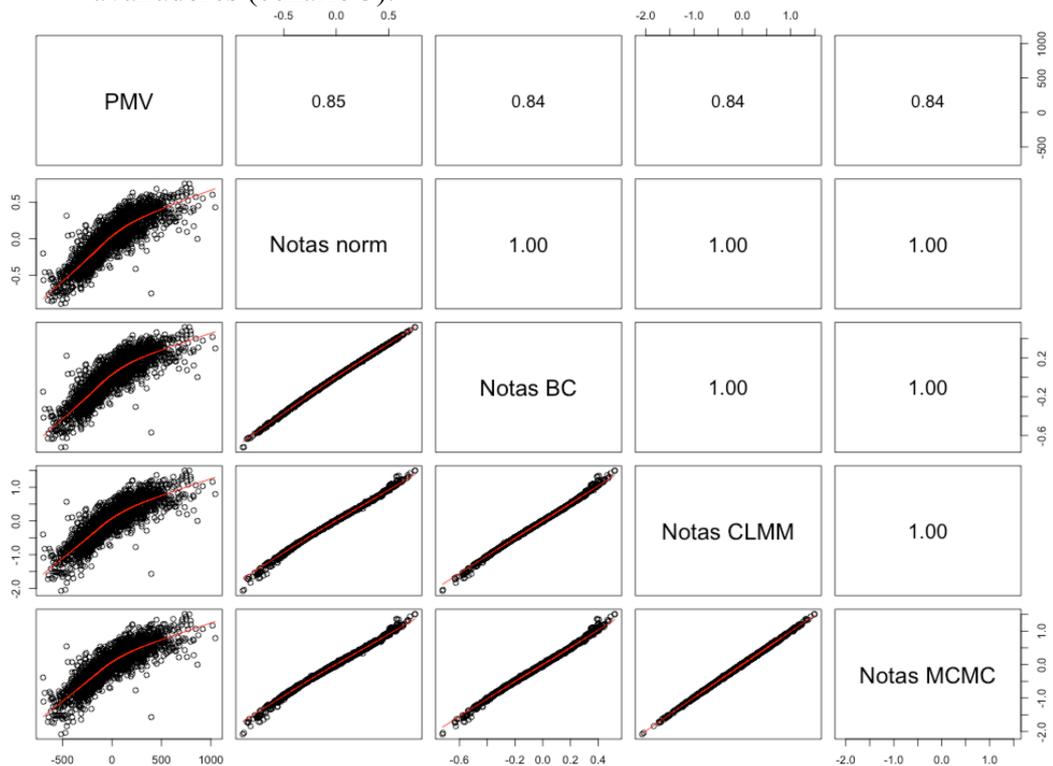
As coincidência dos selecionados entre as diferentes alternativas de análise das notas variaram de 73% a 100% nos três cenários (TABELA 2). Verificou-se uma eficiência de seleção média para PMV com base na nota de vigor de 40%, fato realçado pelos gráficos de dispersão (FIGURAS 2, 3 e 4).

Figura 3 – Painel das correlações de Spearman entre BLUPs dos genótipos pra a características PMV e nota de vigor nas diferentes alternativas considerando três avaliadores (cenário 2).



Fonte: próprio autor.

Figura 4 – Painel das correlações de Spearman entre BLUPs dos genótipos pra a características PMV e nota de vigor nas diferentes alternativas considerando seis avaliadores (cenário 3).



Fonte: próprio autor.

Tabela 2 – Comparativo da eficiência da seleção entre as alternativas de análise das notas de vigor e suas correspondências comparadas com a PMV nos três cenários pela seleção dos 5% melhores.

Eficiência da seleção				
Cenário 1				
	Notas norm	Notas BC	Notas CLMM	Notas MCMC
PMV	0,40	0,44	0,40	0,40
Notas norm		0,81	1,00	0,97
Notas BC			0,81	0,82
Notas CLMM				0,97
Cenário 2				
	Notas norm	Notas BC	Notas CLMM	Notas MCMC
PMV	0,34	0,41	0,41	0,41
Notas norm		0,73	0,78	0,78
Notas BC			0,95	0,95
Notas CLMM				0,98
Cenário 3				
	Notas norm	Notas BC	Notas CLMM	Notas MCMC
PMV	0,42	0,40	0,42	0,42
Notas norm		0,92	0,99	0,98
Notas BC			0,92	0,92
Notas CLMM				0,97

A média de PMV da população original foi de 1277,1g/planta, assim detectou-se que a seleção direta dos 5% melhores resultou em uma estimativa de ganho de quase 50% (TABELA 3). Em contrapartida, a estimativa de ganho para PMV com a seleção indireta mostrou-se inferior, apresentando um ganho médio de 35% nos diferentes cenários. Em geral, ganhos com a seleção mais elevados foram obtidos pelas alternativas MLMG e MLMGB.

Quando a seleção foi realizada com base nos BLUPs dos genótipos, cujos intervalos de predição se encontravam acima da média geral (TABELA 4), procedimento ilustrado na figura 5, constatou-se uma estimativa de ganho direto de 56% para PMV. Enquanto que os ganhos indiretos com base na nota de vigor foram de, aproximadamente, 33%, 47% e 50%, nos respectivos cenários 1, 2 e 3 (TABELA 3).

Tabela 3 – Estimativas do ganho com a seleção dos 5% melhores genótipos para a PMV e notas de vigor para as quatro alternativas de análise e três cenários.

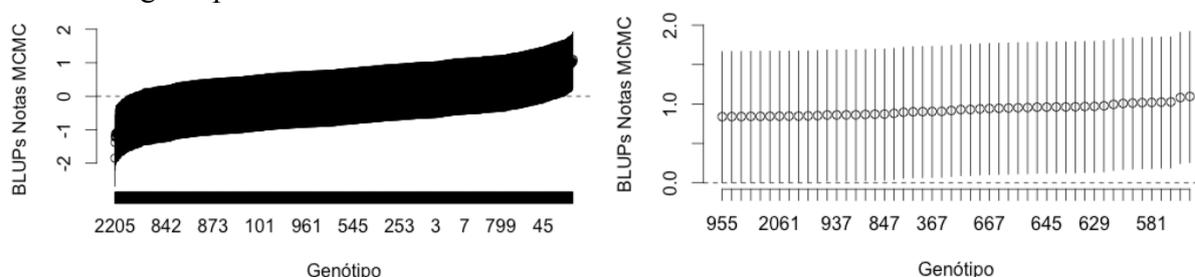
Ganho com a Seleção (g/planta)			
PMV	603.38		
	Cenário 1	Cenário 2	Cenário 3
Notas norm	418.59	423.47	444.38
Notas BC	431.12	444.36	435.75
Notas CLMM	418.59	451.78	442.17
Notas MCMC	421.88	453.62	442.43

Tabela 4 – Estimativas do ganho com a seleção dos genótipos estatisticamente superiores a media geral para a PMV e notas de vigor para as quatro alternativas de análise e três cenários.

Ganho com a Seleção (g/plantas)			
PMV	690,43 (61)		
	Cenário 1	Cenário 2	Cenário 3
Notas norm	515,53 (7)	-	635,91 (18)
Notas BC	278,84 (399)	354,84 (276)	-
Notas CLMM	-	-	-
Notas MCMC	399,39 (50)	576,10 (36)	610,61 (22)

(): indica o número de genótipos selecionados.

Figura 5 Seleção dos BLUPs dos genótipos com intervalos de credibilidade acima da média geral para o cenário 1.



Fonte: próprio autor.

4. DISCUSSÃO

A utilização das notas no melhoramento de espécies forrageiras se justifica devido a sua eficiência em promover a seleção indireta de genótipos superiores. Silva et al. (2013) relataram o sucesso do emprego de notas em *U. ruziziensis* para selecionar genótipos resistentes à *Collaria oleosa*. Resultados semelhantes retratando a eficiência da análise visual foram apresentados por Souza Sobrinho et al. (2009) ao estudar a resistência à cigarrinha. Riday (2009) reportou a eficiência da nota para selecionar genótipos de *red clover* mais produtivos. Além da sua eficiência para selecionar genótipos superiores de forma prática, o

emprego das notas também é utilizada para auxiliar no descarte de genótipos indesejáveis (ATKINS, 1964; BURTON, 1982).

É importante entender que a opção por apresentar as alternativas para análise de notas em cenários distintos se fundamentou na tentativa de representar a realidade dos programas de melhoramento. Em fases mais avançadas de um programa, é comum a utilização de maior número de avaliadores, visto que, nesse momento, é necessário um rigor maior para detectar diferenças reais entre os genótipos. Entretanto, nas fases iniciais, as quais apresentam um grande número de genótipos, a adoção de apenas um avaliador se mostra razoável devido a existência de grande variabilidade genética, o que dispensa o grau de rigor adotado nas fases avançadas.

No presente estudo foram avaliados genótipos na fase inicial de um ciclo de melhoramento em *U. ruziziensis*. Observou-se ampla variabilidade genética para as características PMV e notas de vigor, demonstrando a possibilidade de seleção de genótipos superiores. No entanto, as baixas estimativas de herdabilidade denotam o pronunciado efeito do ambiente na expressão das características PMV e notas de vigor, o que pode estar associado ao delineamento experimental empregado, no caso blocos aumentados.

É importante ressaltar que parte da variação da herdabilidade pode ser explicada pelo fato de suas estimativas se encontrarem em escalas diferentes. Por exemplo, o emprego da transformação Box-Cox resultou no incremento das magnitudes da herdabilidade nos cenários 1 e 2. Besbes et al. (1991) também constataram incremento para o referido parâmetro ao analisar dados da produção de ovos em aves. Porém, essa estimativa é questionável quando associada à deviance de seus modelos, cujos valores indicam um pior ajuste, permitindo reforçar que a transformação pode conduzir a resultados espúrios (JAEGER, 2008; BOLKER, 2009; STROUP, 2015).

Considerando as estimativas da eficiência de seleção (TABELA 3) juntamente com as correlações (figuras 2, 3 e 4) fica evidente que, quanto maior o número de avaliadores, maior é a segurança para efetuar a seleção de genótipos de *U. ruziziensis* de forma indireta para a produção de biomassa, independentemente da alternativa empregada. No caso dos MLM, ao se aumentar o número de avaliadores tem-se uma melhor aproximação da normal pelo teorema do limite central (MOOD; GRAYBILL; BOES, 1974) e, conseqüentemente, melhora o desempenho dos testes estatísticos proporcionando predições mais confiáveis dos valores dos genótipos.

No entanto, ao lidar com dados que não seguem a distribuição normal, é comum o pesquisador lançar mão de transformações, a exemplo da Box-Cox. Observou-se, no presente

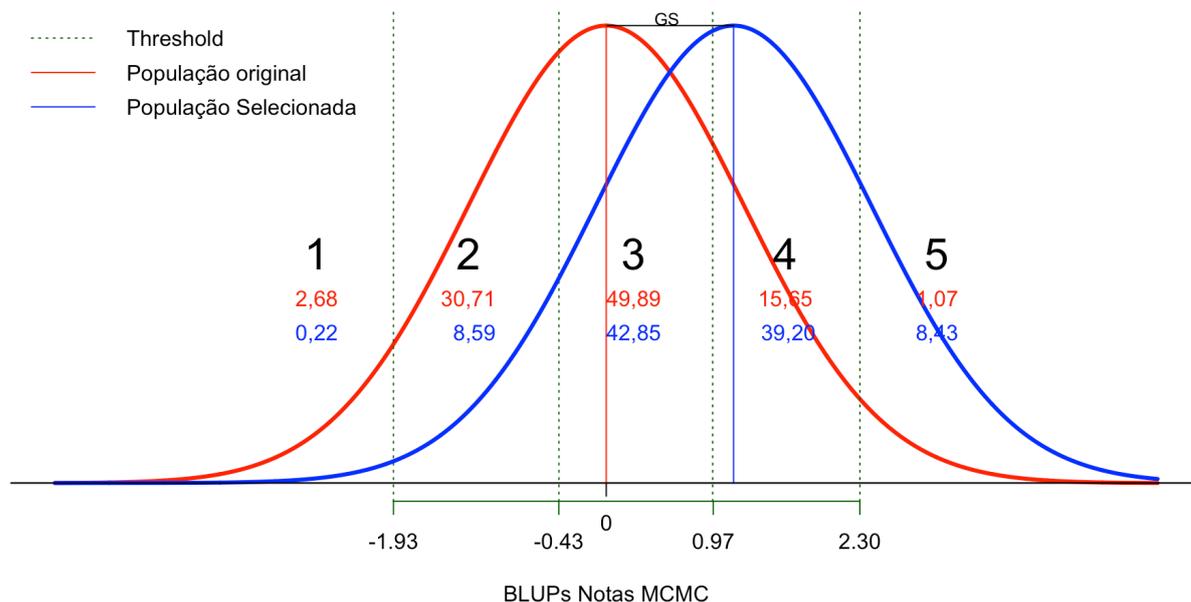
trabalho, que a transformação deve ser utilizada com cautela, pois o atendimento ao pressuposto de normalidade, em alguns casos, não foi assegurado. Verificou-se certa discretização dos BLUPs no cenário 1 (FIGURA 2), sendo possível distinguir grupos distintos, o que corresponde a mesma quantidade de categorias (notas de vigor) adotadas no experimento, reforçando a necessidade da adoção de uma distribuição adequada para a natureza do dado (JAEGER, 2008; BOLKER, 2009; STROUP, 2015).

O ajuste das notas por MLMG é estatisticamente melhor justificado uma vez que considera a natureza da variável aleatória, ou seja, categórica ordinal (AGRESTI, 2007). Seu emprego permite a estimação de limiares das categorias (threshold) (AGRESTI, 2002; GIANOLA; FERNANDO, 1986), viabilizando maior rigor na classificação dos genótipos por considerar seus méritos genéticos classificatórios (FALCONER; MACKAY, 1996). A utilização dos MLMG no presente trabalho permitiu a estimação dos threshold decorrente das notas de vigor atribuídas aos genótipos de *U. ruziziensis* (FIGURA 6). Foi possível notar uma maior concentração de genótipos nas categorias intermediárias, independentemente do gráfico. Contudo, o gráfico vermelho (população original) apresentou uma concentração relativamente maior de genótipos nas categorias inferiores, enquanto no gráfico azul (população melhorada), verificou-se que essa concentração se deu nas categorias 4 e 5.

A tabela 5 expandiu as estimativas dessas proporções de genótipos às categorias para os cenários 2 e 3. Constatou-se que não houve diferenças marcantes entre a utilização de 3 ou 6 avaliadores. Contudo, verificou-se um acréscimo significativo na proporção de genótipos nas categorias superiores pela adoção de um número maior de avaliadores, indicando que o número de avaliadores pode impactar positivamente no ganho com a seleção.

Dentre os critérios comparativos empregados nesse trabalho, aqueles que comparam as alternativas de análise de notas com a PMV retrataram a eficiência da seleção indireta (FIGURAS 2, 3 e 4). Riday (2009) ressaltou a eficiência da avaliação visual em *red clover* ao afirmar que o ganho com a seleção indireta foi tão acurada quanto o ganho com a seleção direta.

Figura 6 – Estimativa do incremento da proporção do número de genótipos, em porcentagem, nas categorias superiores decorrente da seleção com base na nota.



Fonte: próprio autor.

Tabela 5 – Estimativa do ganho com a seleção dos genótipos estatisticamente superiores na escala da nota para notas MCMC nos três cenários.

Categoria	Probabilidade das populações (%)		GS	
	Original	Melhorada		
Cenário 1	1	2,68	0,22	-
	2	30,71	8,59	-
	3	49,89	42,85	-
	4	15,65	39,20	23,55
	5	1,07	8,43	7,36
Cenário 2	1	0,05	> 0,00	-
	2	21,30	1,25	-
	3	73,24	55,10	-
	4	5,38	42,18	36,80
	5	0,01	1,48	1,47
Cenário 3	1	0,07	> 0,00	-
	2	23,25	2,30	-
	3	69,45	55,34	-
	4	7,18	40,31	0,33
	5	0,05	2,05	2,00

GS: Estimativa da proporção do ganho com a seleção proveniente do incremento de indivíduos nas categorias 4 e 5.

A avaliação visual, portanto, se mostra uma estratégia de avaliação interessante para ser utilizada nos programas de melhoramento de *U. ruzizensis*. Sua adoção permite reduzir o

tempo gasto e os custos operacionais de fenotipagem com a realização dos cortes e pesagem dos genótipos, especialmente nas fases iniciais dos ciclos de melhoramento.

5. CONCLUSÃO

As alternativas de análise das notas não mostraram diferenças acentuadas na seleção de genótipos de *U. ruziziensis* para produção de massa verde nos cenários com um, três e seis avaliadores. Porém, as abordagens baseadas em modelos lineares mistos generalizados são mais adequados por considerarem a natureza da variável nota (categórica ordinal).

REFERÊNCIAS

- ABADASSI, J., Mass Selection to Improve Grain Yield in a Traditional Tropical Maize Population. **International Journal of Science and Technology**, v. 4, n. 3, 2014.
- ABREU, G. B.; RAMALHO, M. A. P.; TOLEDO, F. H. R. B.; SOUZA, J. C. de. Strategies to improve mass selection in maize. **Maydica**, Bergamo, v. 55, n. 4, p. 219-225, 2010.
- AGRESTI, A. **An Introduction to Categorical Data Analysis** (second ed.), 2007.
- AGRESTI, A. **Categorical data analysis** (second ed.). New York, NY: John Wiley & Sons, 2002
- ATKINS, R. E. Visual Selection for Grain Yield in Barley. **Crop Science**, v. 4, n. 5, p. 494-497, 1964.
- BATES, D. et al. Fitting Linear Mixed-Effects Models Using {lme4}. **Journal of Statistical Software**. 2015.
- BAYES, T.; PRICE, R.; CANTON, J. **An essay towards solving a problem in the doctrine of chances**. C. Davis, Printer to the Royal Society of London, 1763.
- BESBES, B. et al. Box-Cox transformation of egg-production traits of laying hens to improve genetic parameter estimation and breeding evaluation. **Livestock Production Science**, v. 33, n. 3-4, p. 313-326, 1993.
- BOLKER, B. M. et al. Generalized linear mixed models: a practical guide for ecology and evolution. **Trends in ecology & evolution**, v. 24, n. 3, p. 127-135, 2009.
- BOX, G. E. P.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 211-252, 1964.
- BRESLOW, N. E.; CLAYTON, D. G. Approximate inference in generalized linear mixed models. **Journal of the American statistical Association**, v. 88, n. 421, p. 9-25, 1993.
- BRUCKNER, P. L.; RAYMER, P. L.; BURTON, G. W. Recurrent phenotypic selection for forage yield in rye. **Euphytica**, v. 54, n. 1, p. 11-17, 1991.
- BURTON, G. W. Improved recurrent restricted phenotypic selection increases bahiagrass forage yields. **Crop Science**, v. 22, n. 5, p. 1058-1061, 1982.
- BURTON, G. W. Recurrent restricted phenotypic selection increases forage yields of *Pensacola bahiagrass*. **Crop Science**, v. 14, n. 6, p. 831-835, 1974.
- CHRISTENSEN, R. H. B. Ordinal - regression models for ordinal data. **R package version**, v. 22, 2010.
- FALCONER, D. S.; MACKAY, T. F. C. Threshold characters. **Introduction to Quantitative genetics**, p. 299-311, 1996.
- FÈ, D. et al. Genetic and environmental variation in a commercial breeding program of perennial ryegrass. **Crop Science**, v. 55, n. 2, p. 631-640, 2015.
- FIGUEIREDO, U. J.; NUNES, J. A. Rodrigues; VALLE, C. B. do. Estimation of genetic parameters and selection of *Brachiaria humidicola* progenies using a selection index. **Crop Breeding and Applied Biotechnology**, v. 12, n. 4, p. 237-244, 2012.

- GIANOLA, D. Theory and analysis of threshold characters. **Journal of Animal Science**, v. 54, n. 5, p. 1079-1096, 1982.
- GIANOLA, D.; FERNANDO, R.L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v.63, p.217-244, 1986.
- HADFIELD, J. D. et al. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. **Journal of Statistical Software**, v. 33, n. 2, p. 1-22, 2010.
- HAMBLIN, J.; ZIMMERMANN, M. J. de O. Breeding common bean for yield in mixtures. **Plant Breeding Reviews**, volume 4, p. 245-272, 1986.
- HENDERSON, C. R. **Applications of linear models in animal breeding**. 384 p. 1984.
- HENDERSON, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, p. 423-447, 1975.
- JAEGER, T. F. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. **Journal of memory and language**, v. 59, n. 4, p. 434-446, 2008.
- LOPES, F. C. et al. Composição química e digestibilidade ruminal in situ da forragem de quatro espécies do gênero *Brachiaria*. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 62, n. 4, p. 883-888, 2010.
- MCCULLOCH, C. E. et al. **An Introduction to Generalized Linear Mixed Models**. Biometrics Unit Technical Reports: Number BU-1340-MA. 1997.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**, v. 3. 564 p. 1974.
- NDIKUMANA, J. Etude de l'hybridation entre espèces apomictiques et sexuées dans le genre *Brachiaria*. **Ph.D Thesis**. Université Catholique de Louvain, Louvain-la-Neuve, Belgique. 210 p. 1985.
- NGUYEN, H. T.; SLEPER, D. A. Genetic variability of seed yield and reproductive characters in tall fescue. **Crop science**, v. 23, n. 4, p. 621-626, 1983.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v.58, p.545-554, 1971.
- PINHEIRO, J. C.; CHAO, E. C. Efficient laplacian and adaptive gaussian quadrature algorithms for multilevel generalized linear mixed models. **Journal of Computational and Graphical Statistics** 15(1), 58–81, 2006.
- PRICE, D. L.; CASLER, M. D. Divergent selection for secondary traits in upland tetraploid switchgrass and effects on sward biomass yield. **BioEnergy Research**. 2014.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: <http://www.R-project.org/>. (2015).
- RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical (Agricultural Research in the Tropics)**, v. 37, n. 3, p. 182-194, 2007.

- RESENDE, M.D.V. de. Genética biométrica e estatística no melhoramento de plantas perenes. **Embrapa Informação Tecnológica**, Brasília. 975 p. 2002.
- RIDAY, H. Correlations between visual biomass scores and forage yield in space planted red clover (*Trifolium pratense L.*) breeding nurseries. **Euphytica**, v. 170, n. 3, p. 339-345, 2009.
- SAVIDAN, Y. H., JANK, L., COSTA, J. C. G.; VALLE, C. D. Breeding *Panicum maximum* in Brazil. 1. Genetic resources, modes of reproduction and breeding procedures. **Euphytica**, v. 41, n. 1, p. 107-112, 1989.
- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, v. 52, n. 3-4, p. 591-611, 1965.
- SILVA, D. Maria et al. Genetic Variability of *Brachiaria ruziziensis* Clones to *Collaria oleosa* (Hemiptera: Miridae) Based on Leaf Injuries. **American Journal of Plant Sciences**, v. 4, n. 12, p. 2418, 2013.
- SOUZA SOBRINHO, F. de et al. Repetibilidade de características agrônômicas e número de cortes necessários para seleção de *Urochloa ruziziensis*. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 6, p. 579-584, jun. 2010.
- SOUZA SOBRINHO, F.; LEDO, F. J. S.; KOPP, M. M.; PEREIRA, A. V.; SOUZA, F. F. Melhoramento de gramíneas forrageiras na Embrapa Gado de Leite. In: SIMPÓSIO E CONGRESSO DE FORRAGICULTURA E PASTAGENS, 7., 2009, Lavras. **Anais...** Lavras: Universidade Federal de Lavras. p. 98-115, 2009.
- STROUP, W. W. Rethinking the analysis of non-normal data in plant and soil science. **Agronomy Journal**, v. 107, n. 2, p. 811-827, 2015.
- STUTHMAN, D. D.; STEIDL, R. P. Observed gain from visual selection for yield in diverse oat populations. **Crop Science**, v. 16, n. 2, p. 262-264, 1976.
- TIMBO, A. L. de O. et al. Obtaining tetraploid plants of ruzigrass (*Brachiaria ruziziensis*). **R. Bras. Zootec.**, Viçosa, v. 43, n. 3, p. 127-131, Mar. 2014.
- VALLE, C. B.; JANK, L.; RESENDE, R. M. S. O melhoramento de forrageiras tropicais no Brasil. **Revista Ceres**, v. 56, n. 4, p. 460-472, 2009.
- VALLE, C.B. do; EUCLIDES, V. P. B.; MACEDO, M. C. M. Características das plantas forrageiras do gênero *Brachiaria*. SIMPÓSIO SOBRE MANEJO DA PASTAGEM, v. 17, p. 65-108, 2000.
- WILKINSON, J. M.; HILL, J. Effect on yield and dry-matter distribution of the stay-green characteristic in cultivars of forage maize grown in England. **Grass and Forage Science**, v. 58, n. 3, p. 258-264, 2003.

APÊNDICE A – Rotina utilizada para ajustar a PMV e as alternativas de análise das notas no cenário 2 e para estimar os critérios de comparação.

```

rm(list=ls());ls() # Remove objetos

setwd("//Users//Jfonseca//Desktop//Mestrado//Dissertação Mestrado//Dados") # diretório

dados <- read.table("Dissertação.txt", header = T)
head(dados)
str(dados)
dados$bloco <- factor(dados$bloco)
dados$trat <- factor(dados$trat)

library(lme4) # pacote para MLM

fit.PMV <- lmer(pmv ~ (1|bloco) + (1|trat), data=dados) # ajustar PMV via MLM
fit.PMV0 <- lmer(pmv ~ (1|bloco), data=dados)
summary(fit.PMV)

anova(fit.PMV,fit.PMV0) # LRT
varcomp <- VarCorr(fit.PMV) # extrair BLUPs

## obtendo efeitos aleatórios ##
blup <- ranef(fit.PMV, condVar = TRUE)
blupPMV <- blup$trat ### extrai BLUPs (ef. aleatorio) e variancia
blupPMV.PEV <- attr(blupPMV, "postVar") # atribuir PEV [t=0,975]
blupPMV.SEP <- sqrt(blupPMV.PEV[,1:length(blupPMV.PEV)])

# BLUPs and SEP
tablup <- (cbind(blupPMV,matrix(blupPMV.PEV),blupPMV.SEP))
colnames(tablup) <- c("BLUP","PEV","SEP")
head(tablup)
PEVmedia.pmv <- mean(tablup[1:2204,2])

VG.pmv <- attr(varcomp$trat, "stddev")^2
VB.pmv <- attr(varcomp$bloco, "stddev")^2
VE.pmv <- attr(varcomp, "sc")^2

acuracia.pmv <- sqrt(1-(PEVmedia.pmv/VG.pmv))
acuracia.pmv

h2.pmv <- VG.pmv/(VG.pmv+VB.pmv+VE.pmv)
h2.pmv

```

```
#####
## plotando performance dos BLUPS:## Caterpillar plot

int.conf <- unlist(blupPMV) + qnorm(0.975) * blupPMV.SEP %o% c(-1,1)
genot.ord <- order(unlist(blupPMV))
ic.genot <- int.conf[genot.ord,]
plot(1:2206, unlist(blupPMV)[genot.ord], axes=FALSE, ylim=range(ic.genot),
     xlab="Genótipos", ylab="BLUPs Notas")
axis(1, at=1:2206, labels = genot.ord)
axis(2)
for(i in 1:2206) segments(i, ic.genot[i,1], i, ic.genot[i, 2])
abline(h = 0, lty=2)

#####
#### Seleção de genótipos acima da média #####

blupPMV.ord <- unlist(blupPMV)[genot.ord]
indica <- which(ic.genot[, 1]>0)
blupPMV.ord[indica]

plot(1:length(indica), blupPMV.ord[indica], axes=FALSE, ylim=c(0,1500),
     xlab="Genótipos", ylab="BLUPs Notas")

axis(1, at=1:length(indica), labels = genot.ord[indica])
axis(2)
for(i in 1:length(indica)) segments(i, ic.genot[indica[i],1], i, ic.genot[indica[i], 2])
abline(h = 0, lty=2)

#####
##### Notas #####

dados$nota <- (dados$A+dados$B+dados$C)/3 # notas média de 3 avaliadores
fit.notas <- lmer(nota ~ (1|bloco) + (1|trat), data=dados)
fit.notas0 <- lmer(nota ~ (1|bloco), data=dados)

#####
##### Notas Box-Cox #####

### ajuste do modelo linear fixo para achar lambda (boxcox não calcula lambda para modelo
misto)
fit.lambda <- lm(nota ~ bloco + trat, data=dados)
library(MASS) # pacote para calcular lambda
graf <- boxcox(fit.lambda,
              lambda = seq(-2, 6, length = 10)) # cria graf lambda x logLik
graf # valores lambda x logLik
lambda <- graf$x[which(graf$y==max(graf$y))] # variável transformadora
dados$nota.bc <- (((dados$nota)^lambda)-1)/lambda #cria coluna com notas transformadas

### ajustar modelo linear misto para nota.bc ##
fit.nbc <- lmer(nota.bc ~ (1|bloco) + (1|trat), data=dados)
```

```

fit.nbc0 <- lmer(nota.bc ~ (1|bloco), data=dados)

# teste de normalidade dos resíduos
qqnorm(resid(fit.nbc), main = "", xlab = "", ylab = "")
shapiro.test(resid(fit.nbc))
text(2,-20, "Shapiro-Wilk = 0,91865", cex = 1.6)
text(2.27,-27, "p-valor < 2,2e-16", cex = 1.6)

#####
##### Notas Probit ##### com efeito de 3 avaliadores avaliador
#####

library("ordinal") # Pacote para ajustar MLMG

dados1$score <- factor(dados1$score, ordered=T) # NOTA DEVE ESTAR ORDENADA

### ajustar modelo misto cumulativo linear (MLMG)
fit.np <- clmm(score ~ avaliador + (1|bloco) + (1|parcela) + (1|trat), data=dados1,
              link="probit") # função de ligação probit

#####
##### seleção acima da média ##### constrói gráfico com thresholds

x <- seq(-4,4,length=2000)
y <- dnorm(x,mean=0, sd=1)
plot(x,y, type="l", lwd=2, col = "red", axes=FALSE,
      xlab="BLUPs CLMM", ylab="")
plot.window(xlim=c(-5,5), ylim=c(0,2))
axis(1, pos = 0, c(seq(from=-7, to=7, by=7))) # pos=-0.0795 para igualar eixo
eixo <- c(coef(fit.np)[1],coef(fit.np)[2],coef(fit.np)[3],coef(fit.np)[4])
axis(1, eixo, labels = attr(eixo, "names"), tick=T, col="dark green")
abline(v = coef(fit.np)[1], lty=3, col="dark green") # Atribui estimativa do threshold no graf.
abline(v = coef(fit.np)[2], lty=3, col="dark green")
abline(v = coef(fit.np)[3], lty=3, col="dark green")
abline(v = coef(fit.np)[4], lty=3, col="dark green")
par(new=TRUE)
plot.window(xlim=c(-5,5), ylim=c(0,2))
z <- seq(-4,4,length=2000)
w <- dnorm(x,mean=mean(blupNP.ord[indica]), sd=1)
plot(z,w, type="l", lwd=2, axes=FALSE, col="blue",
      xlab="BLUPs CLMM", ylab="")
segments((mean(blupNP.ord[indica])),0,(mean(blupNP.ord[indica])),max(w),lty=1,col="blue
")
segments(0,(max(y)),(mean(blupNP.ord[indica])),max(w), lty = 1)
segments(0,(max(y)),0,0, lty = 1, col="red")
abline(h = 0, lty = 1, col = "black")

```

```
#####
##### Notas MCMCglimm #####
#####

dados1$score <- as.numeric(dados1$score) # NOTA DEVE SER NUMÉRICA
library("MCMCglimm") # Pacote para ajustar modelos sob enfoque bayesianos

### ajustar modelo misto ordinal no enfoque bayesiano ### (GLMMB)
prior <- list(R=list(V=1, fix=1), Especificação da variância probit
             G=list(
               G1=list(V=1, nu=0.002),
               G2=list(V=1, nu=0.002),
               G3=list(V=1, nu=0.002))) # Especificação da priori gama inversa
fit.MCMC <- MCMCglimm(
  fixed = score ~ avaliador,
  random = ~ trat + bloco + parcela,
  prior = prior,
  family = c("threshold"), nitt = 500000, burnin = 100000,
  thin = 200, data = dados1, pr=T) Especificação do modelo

#### Extrair resultados da análise
DIC <- fit.MCMC$DIC
Deviance <- fit.MCMC$Deviance
VG <- fit.MCMC$VCV[,1]
VB <- fit.MCMC$VCV[,2]
VP <- fit.MCMC$VCV[,3]
VE <- fit.MCMC$VCV[,2]+fit.MCMC$VCV[,3]+1
h2 <- VG/(VG+VE)
Acuracia.MCMC <- sqrt(h2)
#solucao <- fit.model$Sol
results <- as.matrix(cbind(DIC,Deviance,VG,VB,VP,VE,h2,Acuracia.MCMC))
post <- colMeans(results)

MCMCblupN <- fit.MCMC$Sol[,4:2209] # extrair BLUPs

##### Atribuir Cut-Points #####
CP1 <- colMeans(fit.MCMC$CP)[1]
CP2 <- colMeans(fit.MCMC$CP)[2]
CP3 <- colMeans(fit.MCMC$CP)[3]

#### Calcular probabilidades das categorias #####
l <- colMeans(as.matrix(fit.MCMC$Sol[,1])) # variável latente na escala probit
pnorm(-l)
pnorm(CP1-l)-pnorm(-l)
pnorm(CP2-l)-pnorm(CP1-l)
pnorm(CP3-l)-pnorm(CP2-l)
1-pnorm(CP3-l)
```

```
#####
#### Correlação ####
#####

df <- cbind(blupPMV,blupN,blupNBC,blupNP,blupMCMC)
colnames(df) <- c("PMV","Notas norm","Notas BC", "Notas CLMM","Notas MCMC")
matcorS <- cor(df, method = "spearman")

#### Painel dos gráfico #####
panel.cor <- function(x, y, digits = 2, prefix = "", cex.cor)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r <- abs(cor(x, y))
  txt <- format(c(r, 0.123456789), digits = digits)[1]
  txt <- paste0(prefix, txt)
  if(missing(cex.cor)) cex.cor <- 0.6/strwidth(txt)
  text(0.5, 0.5, txt, cex = cex.cor * r)
}
grafcor <- pairs(df, lower.panel = panel.smooth, upper.panel = panel.cor,
  cex.labels = 2, font.labels = 1)

#####
#### Índice de coincidência ####
#####

#função para índice de coincidência
#x: matriz com genótipos ordenados do método 1
#y: matriz com genótipos ordenados do método 2
#s: percentual a ser selecionado

x<-matrix(order(-unlist(blupPMV)))
y<-matrix(order(-unlist(blupMCMC)))

CI=function(x,y,s){
  #x<-x[order(-x[,1]),] #ordena x do maior para o menor
  #y<-y[order(-y[,1]),] #ordena y do maior para o menor
  size<-ceiling(nrow(x)*s) #calcula o tamanho da amostra, n*s
  both<-sum(x[1:size,]%in%y[1:size,]) #verifica quantos dentre o percentual selecionado são
  os mesmos nos 2 métodos
  random<-size*s #calcula o percentual de indivíduos que podem ser coincidentes devido ao
  acaso
  ci<-(both-random)/(size-random) # expressão que calcula o índice de coincidência
  return(ci)
}
CI(x, y,s=0.05) # Calcula eficiência de seleção
```

```
##### GS para PMV #####
z <- cbind(1:2206,blupPMV)
z <- z[order(-z[,2]),]
porcentagem <- ceiling(nrow(z)*0.05) # Calcula quantidade de indivíduos selecionados
mean(z[1:porcentagem,2])

##### Conversão do GS #####
z <- cbind(1:2206,blupPMV)
x <- cbind(1:2206,blupMCMC)
z <- z[order(-z[,2]),]
x <- x[order(-x[,2]),]
porcentagem <- ceiling(nrow(x)*0.05) # Calcula quantidade de indivíduos selecionados
correspondencia <- which((z[,1])%in%(matrix(x[1:porcentagem,1])))
conversao <- z[correspondencia,]
mean(conversao[,2]) # Ganho com a seleção em gramas, pela nota
```