



JOÃO MARCOS NOVAIS

**REAMOSTRAGEM DE UM PLANO EXPERIMENTAL
MULTIAMBIENTES NA SELEÇÃO DE PROGÊNIES DE SOJA**

**LAVRAS – MG
2025**

JOÃO MARCOS NOVAIS

**REAMOSTRAGEM DE UM PLANO EXPERIMENTAL MULTIAMBIENTES NA
SELEÇÃO DE PROGÊNIES DE SOJA**

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

Prof. Dr. José Airton Rodrigues Nunes
Orientador

Dr. Rafael Ravaneli Chagas
Coorientador

**LAVRAS – MG
2025**

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

Novais, João Novais.

Reamostragem de um plano experimental multiambientes na
seleção de progênies de soja / João Novais Novais. - 2025.
44 p.

Orientador(a): José Airton Rodrigues Nunes.

Coorientador(a): Rafael Ravaneli Chagas.

Dissertação (mestrado profissional) - Universidade Federal de
Lavras, 2025.

Bibliografia.

1. *Glycine max* (L.) Merr 2. Delineamento Experimental.
3. Desbalanceamento 4. Eficiência de Seleção. I. Nunes, José
Airton Rodrigues. II. Chagas, Rafael Ravaneli. III. Título.

JOÃO MARCOS NOVAIS

**REAMOSTRAGEM DE UM PLANO EXPERIMENTAL MULTIAMBIENTES NA
SELEÇÃO DE PROGÊNIES DE SOJA**

**RESAMPLING OF A MULTI-ENVIRONMENT EXPERIMENTAL DESIGN FOR
SELECTION OF SOYBEAN PROGENIES**

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

APROVADA em 26 de fevereiro de 2025.

Prof. Dr. José Airton Rodrigues Nunes UFLA

Prof. Dr. Julio Silvio de Souza Bueno Filho UFLA

Dr. Gabriel Mendes Villela GDM SEEDS

Prof. Dr. José Airton Rodrigues Nunes
Orientador

Dr. Rafael Ravaneli Chagas
Coorientador

**LAVRAS – MG
2025**

RESUMO

O delineamento experimental desempenha um papel crucial no planejamento de experimentos em programas de melhoramento de plantas. Neste estudo objetivou-se avaliar o efeito do desbalanceamento de progênies em ensaios multiambientais no delineamento em blocos aumentados para fins de seleção em programa de melhoramento de soja. Os experimentos foram conduzidos na safra 2021/22 em oito localidades nas regiões norte e nordeste do Brasil. A partir dos dados balanceados, foram simulados cenários com diferentes níveis de desbalanceamento, variando de 12,5% a 87,5%. Os dados de todos os cenários foram analisados usando a abordagem de modelos mistos usando o preditor BLUP (*Best Linear Unbiased Predictor*), com (ABLUP) e sem (IBLUP) matriz de parentesco. Para estimação dos componentes de variância genética e ambiental foi usado o método REML (*Residual Maximum Likelihood*). A eficiência dos planos experimentais a partir das estimativas dos parâmetros acurácia seletiva na média de progênie, correlação classificatória de Spearman e índice de coincidência de seleção. O desbalanceamento resultou em planos experimentais otimizados e mais eficientes na alocação de recursos. A análise ABLUP foi superior à IBLUP em todos os cenários. Observou-se uma redução progressiva na acurácia seletiva, na correlação classificatória de Spearman e no índice de coincidência à medida que o desbalanceamento aumentou. Cenários intermediários de desbalanceamento mostraram-se uma opção interessante para a otimização de planos experimentais, pois apresentaram baixa redução nas estimativas dos parâmetros avaliados e elevada eficiência no uso de recursos, especialmente nas fases iniciais de um programa de melhoramento, quando há muitos genótipos e pouca disponibilidade de material genético para teste.

Palavras-chave: *Glycine max* (L.) Merr; delineamento experimental; desbalanceamento; eficiência de Seleção.

ABSTRACT

The experimental design plays a crucial role in planning experiments for plant breeding programs. This study aimed to evaluate the effect of progeny's unbalancing in multi-environment trials using the augmented block design for selection purposes in a soybean breeding program. The experiments were conducted during the 2021/22 growing season in eight locations across the northern and northeastern regions of Brazil. Based on balanced data, scenarios with different levels of unbalancing, ranging from 12.5% to 87.5%, were simulated. The data from all scenarios were analyzed using the mixed model approach with the BLUP (Best Linear Unbiased Predictor), with (ABLUP) and without (IBLUP) a pedigree matrix. The estimation of genetic and environmental variance components was performed using the REML (Residual Maximum Likelihood) method. The efficiency of the experimental designs was assessed based on the estimates of selective accuracy for the progeny mean, Spearman's rank correlation, and the selection coincidence index. The unbalancing resulted in optimized and more efficient experimental designs for resource allocation. The ABLUP analysis outperformed IBLUP in all scenarios. A progressive reduction in selective accuracy, Spearman's rank correlation, and selection coincidence index was observed as imbalance increased. Intermediate imbalance scenarios proved to be an interesting option for optimizing experimental designs, as they exhibited a low reduction in the evaluated parameter estimates and high resource-use efficiency, especially in the early stages of a breeding program when there are many genotypes and limited genetic material available for testing.

Keywords: *Glycine max* (L.) Merr; experimental design; sparse-testing design; selection efficiency.

INDICADORES DE IMPACTO

O uso de delineamentos experimentais desbalanceados pode gerar impactos significativos no setor agrícola e econômico do Brasil. Ao reduzir a área experimental necessária para testar o mesmo número de genótipos sem comprometer a acurácia seletiva, esse método oferece benefícios tanto para empresas públicas quanto privadas em seus programas de melhoramento de soja. Com a implementação de delineamentos desbalanceados, as empresas têm a oportunidade de expandir sua rede de ensaios, explorando uma gama mais ampla de ambientes. Isso permite uma melhor compreensão da interação genótipo por ambiente (GxE), facilitando a identificação de genótipos adaptados e com boa estabilidade. Além disso, os impactos no orçamento necessário para a implantação do ensaio são significativos, possibilitando que recursos financeiros sejam realocados para outras áreas de pesquisa e desenvolvimento. Este planejamento experimental não apenas promove eficiência e economia nos programas de pesquisa de soja, mas também pode ter um impacto social significativo ao incentivar a expansão da rede de ensaios, incluindo regiões com baixo desenvolvimento social, principalmente nas regiões norte e nordeste do país. Ao levar atividades de pesquisa para essas áreas, as empresas podem contribuir para a geração de empregos locais e o aumento da renda para a população residente. Além disso, a expansão da pesquisa em soja nessas regiões não só proporciona oportunidades de emprego direto na condução dos experimentos, mas também estimula o crescimento de setores auxiliares, como serviços de suporte e transporte, impulsionando ainda mais a economia local. A disponibilidade de variedades de soja adaptadas e estáveis pode ter um impacto significativo na produção nacional. A soja é uma das principais commodities do Brasil, e o aumento da produtividade desse cultivo pode ter efeitos positivos em toda a cadeia de suprimentos, desde os produtores até os exportadores. Com a oferta de variedades mais produtivas, o país pode consolidar sua posição como um dos principais exportadores mundiais de soja, gerando receitas adicionais e fortalecendo sua posição no mercado global de commodities agrícolas.

IMPACT INDICATORS

The use of unbalanced experimental designs can have significant impacts on Brazil's agricultural and economic sectors. By reducing the experimental area needed to test the same number of genotypes without compromising selective accuracy, this approach benefits both public and private companies in their soybean breeding programs. Implementing unbalanced designs allows companies to expand their trial networks and explore a broader range of environments. This, in turn, enhances the understanding of genotype-by-environment (GxE) interactions, facilitating the identification of well-adapted and stable genotypes. Furthermore, the financial impact of adopting such designs is substantial, as it enables companies to allocate resources to other areas of research and development. This experimental approach not only improves efficiency and reduces costs in soybean research programs but also has a meaningful social impact by promoting trial expansion into underdeveloped regions, particularly in the northern and northeastern parts of the country. By conducting research in these areas, companies can contribute to local job creation and increased income for local communities. Expanding soybean research in these regions not only generates direct employment opportunities for trial implementation but also stimulates the growth of supporting industries such as logistics and agricultural services, further driving local economic development. Additionally, the availability of well-adapted and stable soybean varieties can significantly impact national production. Soybeans are one of Brazil's most important commodities, and improving crop productivity can create positive ripple effects throughout the supply chain, from farmers to exporters. By offering higher-yielding varieties, Brazil can solidify its position as a leading global soybean exporter, generating additional revenue and strengthening its presence in the international agricultural commodities market.

LISTA DE ILUSTRAÇÕES

- Figura 1** - Planos experimentais de alocação de dez testemunhas (células de cor azul) e dez progênies sob teste (células de cor verde) de soja em oito locais no delineamento em blocos aumentados balanceado (A) e desbalanceado (B)....23
- Figura 2** - Correlação classificatória de Spearman para produtividade de soja em oito cenários de desbalanceamento, comparando os métodos ABLUP e IBLUP ...28
- Figura 3** - Índice de coincidência na seleção dos 10% melhores genótipos para sete cenários de desbalanceamento comparados ao cenário 0 (balanceado)30

LISTA DE TABELAS

Tabela 1 - Descrição dos locais e suas coordenadas geográficas (em decimais)	22
Tabela 2 - Descrição dos oito cenários de desbalanceamento de progênies sob teste de soja em ensaios multiambientais	23
Tabela 3 - Detalhamento do número de progênies, testemunhas, locais e parcelas para cada cenário de desbalanceamento	26
Tabela 4 - Estimativas de parâmetros genéticos e ambientais para produtividade referente à avaliação de progênies de soja sob diferentes cenários de desbalanceamento em ensaios multiambientais no delineamento em blocos aumentados	27

SUMÁRIO

1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	13
2.1	Melhoramento genético de soja.....	13
2.2	Interação genótipo por ambiente em soja.....	14
2.3	Planos experimentais em ensaios multiambientes	18
3	MATERIAL E MÉTODOS	22
3.1	Locais, genótipos e delineamento experimental	22
3.2	Análise estatística dos dados.....	24
4	RESULTADOS E DISCUSSÃO	26
5	CONCLUSÕES	31
	REFERÊNCIAS.....	32
	APÊNDICE A - PRODUTO TÉCNICO ATRELADO	38

1 INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) é uma leguminosa de grande importância global, classificada como a quarta espécie mais cultivada em todo o mundo. Segundo o Departamento de Agricultura dos Estados Unidos, a produção mundial ultrapassou 350 milhões de toneladas de grãos na safra 2021/22, com o Brasil contribuindo com 35% desse total (USDA, 2022).

Segundo dados da Companhia Nacional de Abastecimento, a produtividade da soja no Brasil apresentou um aumento significativo ao longo dos anos (CONAB, 2022). Na década de 1970, o rendimento médio era de 1.748 kg/ha, e na safra 2021/22, esse valor alcançou 3.526 kg/ha. A produtividade da leguminosa é influenciada por diversos fatores, incluindo condições climáticas, práticas de adubação, cultivares de soja selecionadas e a implementação de tecnologias que apoiam os agricultores nas tomadas de decisão.

O melhoramento de soja desempenha um papel crucial no sucesso do cultivo em diversas regiões do país, impactando positivamente o aumento da produtividade por meio do desenvolvimento de cultivares adaptadas e tolerantes a estresses bióticos e abióticos (EMBRAPA, 2021). Para atender à crescente demanda mundial por essa oleaginosa, empresas públicas e privadas estão cada vez mais direcionando recursos para pesquisas, visando desenvolver e/ou aplicar as melhores ferramentas que auxiliem na seleção das cultivares de soja mais promissoras.

Nas fases iniciais do programa de melhoramento de soja, é comum a baixa disponibilidade de sementes e uma grande quantidade de genótipos a serem testados. Nesse contexto, a escolha do delineamento experimental torna-se crucial, apresentando impactos substanciais em termos financeiros, logísticos e de recursos humanos (MARTIN et al, 2007).

O delineamento em blocos aumentados, proposto por Federer (1956), o *p-rep*, sugerido por Cullis, Smith e Coombes (2006), e mais recentemente, o “*sparse testing design*”, reportado por JARQUIN et al. (2020) e CRESPO-HERRERA L. et al (2021), são exemplos de delineamentos otimizados que se destacam em situações caracterizadas por limitação de área experimental, muitos genótipos a serem testados, restrição de material vegetal e recursos financeiros limitados. O “*sparse testing design*” tem destacada importância na otimização de ensaios multiambientes (MET) em soja, em que os genótipos não são testados em todos os ambientes. MET permitem estimar e explorar a interação genótipos x ambientes e, com isso, selecionar genótipos que se destaquem pela sua estabilidade frente as condições macroambientais. Jarquin et al. (2020) reportam que a otimização de MET usando “*sparse*

testing designs” podem contribuir para uma alocação mais judiciosa de recursos, aumentar a capacidade de avaliação de genótipos e, com isso, aumentar a intensidade de seleção, além de aumentar a acurácia de seleção pela melhor amostragem de ambientes que compõem a população alvo de ambientes.

Ante o exposto, neste trabalho objetivou-se avaliar o efeito do desbalanceamento de progênies em ensaios multiambientais no delineamento em blocos aumentados para fins de seleção em programa de melhoramento de soja. Além disso, comparar diferentes cenários de desbalanceamento mediante estimação dos componentes da variância genética e ambiental, acurácia seletiva, correlação classificatória e índice de coincidência da seleção; e adicionalmente avaliar o efeito da inclusão da matriz de parentescos nos cenários estudados.

2 REFERENCIAL TEÓRICO

2.1 Melhoramento genético de soja

A soja (*Glycine max* (L.) Merrill) pertence à família das Fabaceae. Os primeiros registros de uso da soja pelo ser humano remontam a cerca de 3 mil anos a.C., na região leste da Ásia, mais especificamente na antiga China. A soja ancestral apresentava características fenotípicas bastante distintas das variedades comerciais atuais. Naquela época, a soja era uma planta perene, de porte rasteiro e com uma ramificação pronunciada. Com a domesticação da soja e o progresso no conhecimento em melhoramento genético, houve intensas modificações na arquitetura da soja para atender às necessidades humanas. A soja cultivada hoje é uma planta herbácea, de ciclo anual, com porte ereto e tipo de crescimento que pode ser determinado, semideterminado ou indeterminado (BONATO, E.; BONATO, A., 1987).

No Brasil, o cultivo da soja é registrado desde 1882, mas apenas na década de 70 tornou-se uma cultura de importância econômica. Desde então, foram envidados esforços significativos, tanto por instituições públicas quanto privadas, para aprimorar as características de interesse da soja e torná-la adaptada ao cultivo nas diversas regiões do Brasil (RAMALHO; MARQUES; LEMOS, 2021).

Outro fator que incentivou a criação e expansão de programas de melhoramento da soja foi a promulgação da Lei de Proteção de Cultivares (Lei Nº 9.456, de abril de 1997), uma legislação que estabelece regras e diretrizes para a proteção dos direitos de propriedade intelectual sobre variedades de soja desenvolvidas por melhoristas ou empresas de sementes (RAMALHO; MARQUES; LEMOS, 2021).

Inicialmente, o cultivo da soja estava restrito à região sul do Brasil devido à sua característica fisiológica de determinar o comprimento do ciclo com base no fotoperíodo. As primeiras variedades, quando cultivadas em regiões de baixa latitude, acumulavam rapidamente o número de horas de escuridão necessário para o florescimento, o que resultava em plantas de baixa estatura e baixo rendimento, tornando seu cultivo inviável nessas regiões (ALMEIDA et al., 1999; DALL'AGNOL, 2011).

O grande sucesso na expansão do cultivo de soja no Brasil foi o desenvolvimento, por meio de programas de melhoramento genético, de variedades que apresentavam um período juvenil longo, período no qual a soja ainda não responde ao fotoperíodo. Isso permitiu que as plantas se desenvolvessem principalmente no aspecto vegetativo, resultando em maiores

produtividades (DONG et al, 2012; RAMALHO; MARQUES; LEMOS, 2021; VIANA et al., 2009).

Aumentar a produtividade é a principal característica buscada em um programa de melhoramento, embora características como tolerância a estresses bióticos e abióticos também sejam importantes (MARANNA et al., 2021).

Os estresses bióticos são os maiores responsáveis pela redução de produtividade da soja. As principais doenças que acometem a soja no Brasil incluem a ferrugem asiática (*Phakopsora pachyrhizi*), a mancha-alvo (*Corynespora cassiicola*), a podridão radicular de *Phytophthora* (*Phytophthora sojae*), a bacteriose (*Xanthomonas axonopodis* pv. *glycines*) e os nematóides de galha (*Meloidogyne* spp.) e cisto (*Heterodera glycines*). Essas enfermidades podem causar danos significativos às plantações, tornando essencial o desenvolvimento de cultivares resistentes e a implementação de práticas de manejo eficazes para garantir a sustentabilidade e produtividade da cultura da soja (ALLEN et al., 2017; LIN et al., 2022).

O melhoramento de cultivares de soja também visa desenvolver variedades resilientes às adversidades climáticas que ocorrem durante o ciclo, como estresse hídrico devido à falta ou excesso de chuvas, altas temperaturas, deficiência nutricional no solo e presença de alumínio tóxico. A tolerância da soja ao déficit hídrico destaca-se como o principal foco de estudo por parte de empresas públicas e privadas, dada sua significativa influência na produtividade e sua ocorrência frequente em determinadas regiões de cultivo da cultura (LI, 2019; MUTAVA, 2015; RUPESH, 2014).

Na busca por genótipos altamente produtivos, estáveis e resistentes às principais pragas e doenças, além de serem tolerantes a estresses abióticos, os programas de melhoramento contam com uma rede de experimentos distribuídos em diferentes regiões do país. Essa estratégia busca uma estimativa mais precisa do efeito do fenótipo, permitindo a decomposição dos efeitos genéticos, ambientais e a interação entre genótipo e ambiente.

2.2 Interação genótipo por ambiente em soja

A interação genótipos por ambientes (GxE) é um fenômeno amplamente conhecido no reino animal e vegetal e pode ser entendida como a resposta diferencial do genótipo de acordo com o ambiente no qual está inserido. As plantas possuem informações genéticas que determinam suas características, mas essas características podem ser expressas de maneiras diferentes dependendo do ambiente em que a planta está inserida. Fatores ambientais como luz,

temperatura, umidade, solo e disponibilidade de nutrientes podem modular a expressão dos genes e, assim, influenciar o crescimento, desenvolvimento e desempenho das plantas (RAMALHO et al., 2012).

A interação genótipo x ambiente é um grande complicador na seleção e recomendação de cultivares de soja porque as respostas das plantas às condições ambientais podem variar significativamente entre diferentes locais e períodos de cultivo. Em outras palavras, uma cultivar que se destaca em um determinado ambiente pode não ter o mesmo desempenho em outro devido às influências específicas do ambiente sobre a expressão dos genes (GESTEIRA et al., 2018; SOARES et al., 2017).

A interação genótipo x ambiente pode ser classificada em dois tipos principais: interação simples ou complexa. Na interação simples, ocorre uma mudança na magnitude do desempenho dos genótipos em diferentes ambientes, mas a ordem relativa desses genótipos permanece constante. Por outro lado, na interação complexa, os genótipos respondem de maneira diferenciada aos diferentes ambientes, resultando em uma mudança na classificação desses genótipos. (RAMALHO et al., 2012). A interação complexa é a que exerce maior impacto na tomada de decisão por parte do melhorista e é, de fato, o tipo de interação mais prevalente em experimentos com soja no Brasil (SOARES et al., 2017).

A complexidade na interação GxE destaca a importância de realizar ensaios de cultivares em diferentes locais e condições representativas da diversidade ambiental em que as plantas serão cultivadas. Além disso, a utilização de modelos estatísticos e técnicas de análise que levam em consideração a interação GxE pode aprimorar a precisão das recomendações de cultivares (BIANCHI et al., 2020).

O estudo conduzido por Souza et al. investigou a produtividade de 20 cultivares de soja em duas localidades no sul do Brasil, avaliando-as em três diferentes épocas de plantio. Os resultados revelaram a presença de interação genótipo x ambiente, tanto em relação às distintas épocas de plantio quanto às diferentes localidades. Notavelmente, a análise indicou uma predominância da interação complexa sobre a interação simples. Esse estudo destaca a influência significativa dos fatores ambientais, tanto em termos de variação sazonal quanto de variação geográfica, na resposta das cultivares de soja (SOUZA et al., 2021).

Os efeitos da interação entre genótipo e ambiente podem ser mitigados ao direcionar a avaliação e recomendação de genótipos para áreas com características semelhantes, o que resulta em resultados mais precisos. Ao adaptar a escolha de genótipos às condições específicas

de cada ambiente, é possível explorar o máximo potencial das plantas, tirando proveito dos efeitos da interação entre genótipo e ambiente (YAN, 2016).

A estratificação ambiental pode ser uma alternativa na seleção e recomendação de cultivares para ambientes específicos. (PACHECO et al, 2017). A extensa área de cultivo de soja no Brasil, que abrange diferentes regiões com variabilidade ambiental significativa, destaca a importância de uma abordagem estratificada para compreender e direcionar recomendações de cultivares. A subdivisão em macro e micro-regiões sojícolas permite considerar as particularidades ambientais de cada região, otimizando a adaptação das cultivares às condições locais (KASTER; FARIAS, 2012).

A influência da interação genótipo x ambiente se manifesta nos caracteres qualitativos, mas se torna ainda mais notória nos caracteres quantitativos, revelando a complexidade na seleção de genótipos de soja com alto rendimento em diferentes ambientes (PEREIRA et al., 2019; SQUILASSI, 2003).

Em função da GxE, para seleção e recomendação de cultivares de soja é necessário a adoção de análises de adaptabilidade e estabilidade, uma vez que os fatores edafoclimáticos exercem uma influência preponderante na capacidade de adaptação e estabilidade dos genótipos de soja em termos de produtividade (BORÉM; MIRANDA; NETO FRITSCHÉ, 2021; RAMOS JUNIOR et al., 2017).

A adaptabilidade de um genótipo refere-se à sua habilidade em tirar proveito dos efeitos do ambiente para assegurar um alto rendimento. Por outro lado, a estabilidade está associada à capacidade de manter ou prever a produção em diferentes ambientes (BORÉM; MIRANDA; NETO FRITSCHÉ, 2021). Portanto, um genótipo adaptável se ajusta bem a condições ambientais variáveis, resultando em altos rendimentos, enquanto a estabilidade indica a consistência desse desempenho em ambientes diversos.

Diversas abordagens foram desenvolvidas para estudos da adaptabilidade e estabilidade de genótipos de soja em experimentos multiambientais. (CUCOLOTTO, 2007). Entre os mais utilizados, destacam-se as metodologias AMMI (Additive Main Effect and Multiplicative Interaction) e GGE Biplot (*genotype main effects + genotype environment interaction*) (SILVA; BENIN, 2012; SOUZA et al., 2021).

A AMMI é uma abordagem que combina técnicas de análise de variância com análise de componentes principais. Ela é projetada para separar os efeitos aditivos (principalmente genéticos) dos efeitos multiplicativos (interação genótipo x ambiente). A AMMI produz biplots, gráficos bidimensionais que representam visualmente padrões de interação. Isso ajuda a

identificar genótipos estáveis e adaptáveis em diferentes ambientes. A interpretação desses biplots permite uma compreensão mais clara dos padrões de resposta dos genótipos aos ambientes (GAUCH, 2013; RANI et al., 2023)

A GGE é uma abordagem que incorpora conceitos da estatística multivariada e visa representar a interação GxE por meio de um modelo gráfico. Ela se concentra em descrever os efeitos principais (genótipos e ambientes) e os efeitos de interação GE, fornecendo uma visualização clara das tendências. Os biplots resultantes ajudam a identificar genótipos adaptáveis e ambientes discriminativos. A GGE também utiliza o conceito de "mega-ambientes" para agrupar ambientes similares (CARVALHO et al. 2021; YAN et al, 2007).

Murphy et al. (2009) conduziram um estudo sobre a interação genótipo x ambiente (GxE) para o teor de isoflavonas em variedades de soja, concluindo que os métodos de análise de variância de estabilidade de Shukla (1972) e o coeficiente de regressão de Finlay-Wilkinson (1963) apresentaram concordância e foram eficazes na detecção de diferenças significativas entre os genótipos. As análises gráficas de estabilidade, que também avaliam simultaneamente o desempenho, como a análise média-CV (Francis and Kannenberg's, 1978) e o biplot GGE (YAN et al., 2000), também estavam em concordância e podem ser utilizados em casos de presença de interação GxE (MURPHY et al., 2009).

As diversas ferramentas estatísticas e desenhos experimentais discutidos na literatura para avaliar a interação genótipo x ambiente são comumente recomendados para estágios mais avançados dos programas de melhoramento, nos quais apenas alguns genótipos são avaliados em múltiplos ambientes. No entanto, os melhoristas enfrentam o desafio de compreender e lidar com a GxE nos estágios iniciais do programa, que envolvem a avaliação de muitos genótipos (BIANCHI et al., 2020).

A compreensão da interação genótipo por ambiente pode se tornar mais complexa quando os dados são coletados sem critério e rigor adequados. A utilização de delineamentos experimentais otimizados, aliada ao emprego de técnicas estatísticas apropriadas, é crucial para assegurar uma estimativa precisa dos efeitos dos componentes de variância do caráter em estudo. Essa abordagem é fundamental para garantir o sucesso do programa de melhoramento genético.

2.3 Planos experimentais em ensaios multiambientais

O planejamento do experimento é fundamental para garantir a validade dos resultados de um dado experimento. O pesquisador deve definir cuidadosamente as variáveis dependentes e independentes, o número de testes, e as condições para a realização dos testes (CARPINETTI, 2009).

O delineamento experimental é uma parte crucial do planejamento de experimentos científicos, pois determina como os tratamentos serão atribuídos às unidades experimentais, permitindo a organização e controle das variáveis experimentais. Cada tipo de delineamento tem suas próprias características e objetivos específicos. São exemplos de delineamentos: inteiramente ao acaso, blocos casualizados, quadrado latino, blocos incompletos, reticulados e blocos aumentados (AMABILE; VILELA; PEIXOTO, 2018).

A escolha do delineamento experimental apropriado requer uma consideração cuidadosa de vários critérios essenciais. Primeiramente, é importante definir claramente os objetivos do estudo, diferenciando entre exploração, descrição, explicação ou confirmação, para orientar a seleção do desenho mais adequado. Identificar as variáveis de interesse e a capacidade de controlar variáveis externas são passos fundamentais, assim como avaliar a necessidade de repetições e aleatorização para reforçar a validade estatística. A viabilidade prática, considerando recursos disponíveis, é um aspecto pragmático a ser ponderado, juntamente com a natureza dos dados a serem coletados (CLEWER; SCARISBRICK, 2001).

A determinação do plano experimental a ser adotado requer a consideração de diversos aspectos, incluindo a disponibilidade e qualidade da área, volume de material genético disponível, a constituição genotípica da planta, as características sob estudo e os parâmetros genéticos associados a essas características (MARTIN et al., 2004).

A herdabilidade é um parâmetro genético frequentemente adotado para orientar o plano experimental. Um experimento para características de baixa herdabilidade, como produtividade de grãos, a estratégia de replicação e a distribuição dos tratamentos devem ser cuidadosamente planejadas para maximizar a capacidade de separar os efeitos genéticos dos efeitos ambientais, possibilitando uma interpretação mais precisa dos resultados (RAMALHO; FERREIRA; DE OLIVEIRA, 2012; RESENDE; ALVES, 2022)

A eficiência do delineamento experimental em termos de controle local pode ser estudada por meio do teste F de Snedecor, a variabilidade espacial dentro de estratos ambientais

homogêneos, ou blocos, pode ser examinada por meio do coeficiente de correlação intraclasse entre indivíduos de uma mesma parcela (C2) (STURION; RESENDE, 2005).

Para experimentos em multiambientes, os delineamentos experimentais precisam ser mais robustos para melhor controlar os fatores aleatórios ao experimento. O Delineamento em Blocos Casualizados (DBC) é frequentemente escolhido em experimentos multiambientes devido à sua capacidade de oferecer maior controle sobre variações não controlada ao agrupar unidades experimentais em blocos homogêneos, além de oferecer informações valiosas sobre a interação GxE, fator de grande interesse na área de melhoramento genético de plantas (FREITAS, 2022).

Outro delineamento bastante utilizado em ensaios multiambientes é o delineamento em blocos incompletos, no qual nem todos os tratamentos estão presentes dentro de cada bloco. Este desenho é especialmente indicado quando há a necessidade de testar muitos genótipos em múltiplos ambientes, mas os recursos disponíveis não permitem a implementação do experimento com blocagem e repetições para todos os tratamentos (AKINWALE, 2021).

Existem dois tipos principais de delineamentos incompletos: o delineamento em blocos incompleto balanceado (BIB) e o delineamento em blocos incompletos parcialmente balanceados (PBIB). No BIB, todas as combinações entre dois tratamentos devem ocorrer juntas dentro de cada bloco. Por outro lado, no PBIB, essa exigência não é necessária, sendo possível comparar tratamentos que não ocorrem dentro do mesmo bloco. Entre os tipos de PBIB, o alfa lattice é mais preferido devido à sua praticidade, flexibilidade e versatilidade em comparação com o balanceado (LAWSON, 2015)

Mona et al. (2018) compararam o delineamento em blocos casualizados (DBC) com o desenho alfa lattice para avaliar a produtividade de feijão-fava (*Vicia faba* L.). Concluíram que o alfa lattice foi mais preciso e eficaz na redução do erro experimental em comparação com o DBC, destacando sua capacidade superior de detectar pequenas diferenças significativas entre as médias dos genótipos. O alfa lattice apresentou valores mais baixos de coeficiente de variação (CV) e valor de P, juntamente com valores mais altos de eficiência relativa (RE) e função de poder, evidenciando sua adequação superior ao DBC para o experimento em questão (MONA et al., 2018). Além disso, HOEFLER et al. (2020) evidenciaram que o delineamento alfa lattice foi mais eficaz do que o DBC no controle da heterogeneidade espacial em ensaios com alta e baixa variação espacial.

Os desenhos aumentados (Federer 1956) ou *p-rep* (CULLIS; SMITH; COOMBES, 2006) fazem parte da classe de delineamento incompleto e são frequentemente utilizados nas

etapas iniciais de um programa de melhoramento por lidarem bem com experimentos em que se deseja testar muitos genótipos.

Os desenhos aumentados propostos por Federer (1956) são delineamentos nos quais apenas as testemunhas recebem mais de uma repetição, enquanto os tratamentos regulares são alocados sem repetição. Esse método contorna as limitações comuns em ensaios preliminares, como a escassez de material genético e a avaliação de um grande número de genótipos, ao empregar blocos aumentados (MOEHRING; WILLIAMS; PIEPHO, 2014).

Os delineamentos aumentados podem apresentar uma desvantagem por requerer um grande número de parcelas para alocar testemunhas, que são variedades já bem estabelecidas e de pouco interesse direto para o melhorista (MOEHRING; WILLIAMS; PIEPHO, 2014). Diante dessa consideração, Cullis, Smith e Coombes (2006) apresentaram o desenho *p-rep*, fundamentado na replicação de uma porcentagem (p) das linhas de teste, enquanto as linhagens restantes são alocadas sem replicação nas parcelas remanescentes.

Em um estudo de simulação conduzido por Cullis, Smith e Coombes (2006) foi demonstrado que, para um tamanho de ensaio fixo, os designs "*p-rep*" proporcionaram ganhos genéticos superiores em comparação com os chamados "*grid-plot designs*", que são normalmente empregados em ensaios preliminares de programas de melhoramento.

Peternelli et al. (2009) conduziram uma comparação entre delineamentos aumentados e delineamentos com repetições. Suas conclusões sugerem que, em situações de restrição de recursos, a eficiência de seleção é otimizada ao testar um maior número de genótipos em delineamentos aumentados, em vez de reduzir a quantidade de genótipos para incluir repetições nos ensaios.

Nos estágios iniciais do programa de melhoramento, há muitos genótipos a serem testados. As limitações de recursos e de material genético disponível nessas fases reduzem o número de locais e parcelas que podem ser utilizados para cada tratamento. Dada a herdabilidade relativamente baixa para a característica de rendimento de grãos e a presença de interação genótipo por ambiente, é crucial adotar desenhos experimentais eficientes e métodos de análise apropriados (CULLIS et al. 2020; MOEHRING; WILLIAMS; PIEPHO, 2014).

Além dos delineamentos clássicos citados anteriormente, os desenhos experimentais atuais são conhecidos como "*Optimal model-based*" e refere-se a designs de ensaios de campo que são desenvolvidos com base em modelos matemáticos otimizados para maximizar a eficiência da coleta de dados ou a precisão das análises subsequentes (BAILEY, 2008). O modelo usado no planejamento experimental em programas de melhoramento normalmente é

um modelo linear misto, que também é utilizado nas análises subsequentes (CULLIS et al. 2020).

A maioria dos delineamentos experimentais demanda a alocação de todos os tratamentos regulares em cada ambiente avaliado. Contudo, mesmo com os benefícios otimizados dos desenhos aumentados que eliminam a necessidade de repetição das linhas de teste, essa abordagem não é suficiente para viabilizar economicamente ensaios multiambientais com muitos genótipos em ensaios preliminares. Diante desse desafio, Montesino-López et al. (2023) testaram um novo delineamento denominado "*sparse testing design*", no qual apenas parte das linhagens são avaliadas em ambientes específicos, enquanto outras têm seus valores preditos por meio de modelagem estatística utilizando dados genotípicos via matriz de parentesco (MONTESINOS-LÓPEZ et al., 2023).

A aplicação da matriz de parentesco revela-se extremamente vantajosa, dado que a maioria dos programas de melhoramento já dispõe desse conjunto de informações. Integrar essa matriz no processo de seleção aprimora a eficácia do programa sem acarretar custos significativos (BUTLER; SMITH; CULLIS, 2014).

Atanda et al. (2022) conduziram pesquisas envolvendo ensaios preliminares desbalanceados em trigo e chegaram à conclusão de que a aplicação da genômica na previsão dos valores genéticos de genótipos em locais não testados pode proporcionar benefícios significativos. No entanto, é crucial considerar a sobreposição de genótipos entre os locais, a fim de estabelecer uma conectividade que potencialize a precisão da estimativa da correlação genética entre os ambientes.

O emprego de ensaios desbalanceados, principalmente associados à genômica, pode ser muito benéfico em um programa de melhoramento, permitindo que mais genótipos possam ser avaliados sem aumentar o tamanho da área experimental, ou por reduzir a área necessária testando o mesmo número de genótipos. Este delineamento experimental otimiza o emprego de recursos financeiros e operacionais com um impacto minimamente negativo na precisão de seleção (MONTESINOS-LÓPEZ et al., 2023).

3 MATERIAL E MÉTODOS

3.1 Locais, genótipos e delineamento experimental

Os dados de produtividade de soja (t/ha) foram obtidos do banco de dados da empresa GDM referente aos experimentos conduzidos na safra 2021-2022. Os experimentos foram conduzidos em oito locais situados nos estados do Maranhão, Tocantins, Piauí e Bahia, regiões que compreendem a macrorregião sojícola 5 do Brasil (Tabela 1).

Tabela 1 - Descrição dos locais e suas coordenadas geográficas (em decimais).

Locais	Estado	Latitude	Longitude
Baixa Grande do Ribeiro	Piauí	-8,280528	-45,306126
Uruçuí	Piauí	-7,284753	-44,263549
Tasso Fragoso	Maranhão	-8,404459	-45,971806
Caseara	Tocantins	-9,350122	-49,885794
Santa Rosa do Tocantins	Tocantins	-11,202879	-48,172715
Correntina	Bahia	-13,860562	-46,197318
Porto Nacional	Tocantins	-10,511173	-48,267979
Campos Lindos	Tocantins	-8,203212	-46,57145

Fonte: Do autor (2025)

Foram avaliadas 450 progênies da geração F2:4 de soja provenientes de 81 famílias distintas e mais 10 testemunhas, totalizando 460 genótipos. As testemunhas utilizadas foram NEO820 IPRO, M8220 I2X, TMG2383 IPRO, ST820 I2X, DM82K84RSF CE, DAG 8221 I2X, CZ58B23 I2X, DM82I78 IPRO, BMX DOMÍNIO IPRO e M8349 IPRO. O delineamento experimental empregado foi o de blocos aumentados constituído por nove blocos, cada um contendo 60 parcelas. Em cada bloco foram alocadas aleatoriamente 10 testemunhas e 50 progênies de soja. Cada parcela foi composta por quatro linhas de quatro metros, com espaçamento de 0,5 metro entre as linhas, sendo consideradas úteis as duas linhas centrais. O estudo teve como foco a avaliação da produtividade de grãos, expressa em toneladas por hectare (t/ha).

Para otimizar o plano experimental dos ensaios multiambientais, foram estabelecidos oito níveis de desbalanceamento, variando exclusivamente a distribuição das progênies sob teste entre os ambientes usando o princípio dos *sparse-design models*. Os níveis de desbalanceamento foram de 0% a 87,5%, representando os cenários simulados C0 a C7, respectivamente (Tabela 2). O número de ambientes por genótipo indica quantas observações

cada genótipo obteve, enquanto o número total de ambientes em cada análise permaneceu constante, ou seja, oito ambientes (Tabela 2).

Tabela 2 - Descrição dos oito cenários de desbalanceamento de progênies sob teste de soja em ensaios multiambientais.

Cenário	Número de ambientes por genótipo	Número de ambientes por análise	Desbalanceamento (%)
0	8	8	0
1	7	8	12,5
2	6	8	25
3	5	8	37,5
4	4	8	50
5	3	8	62,5
6	2	8	75
7	1	8	87,5

Fonte: Do autor (2025)

Vale ressaltar que as testemunhas permaneceram balanceadas em todos os cenários, com o mesmo *layout* de parcelas do delineamento balanceado. Cada experimento foi planejado para garantir a amostragem dos oito ambientes em todos os cenários.

Para exemplificar, um plano de ensaios multiambientais no delineamento em blocos aumentados desbalanceado (CENÁRIO 4) é ilustrado na Figura 1.

Figura 1 - Planos experimentais de alocação de dez testemunhas (células de cor azul) e dez progênies sob teste (células de cor verde) de soja em oito locais no delineamento em blocos aumentados balanceado (A) e desbalanceado (B).

Genótipo	local 1	local 2	local 3	local 4	local 5	local 6	local 7	local 8	total
test 1	1	1	1	1	1	1	1	1	8
test 2	1	1	1	1	1	1	1	1	8
test 3	1	1	1	1	1	1	1	1	8
test 4	1	1	1	1	1	1	1	1	8
test 5	1	1	1	1	1	1	1	1	8
test 6	1	1	1	1	1	1	1	1	8
test 7	1	1	1	1	1	1	1	1	8
test 8	1	1	1	1	1	1	1	1	8
test 9	1	1	1	1	1	1	1	1	8
test 10	1	1	1	1	1	1	1	1	8
progênie 1	1	1	1	1	1	1	1	1	8
progênie 2	1	1	1	1	1	1	1	1	8
progênie 3	1	1	1	1	1	1	1	1	8
progênie 4	1	1	1	1	1	1	1	1	8
progênie 5	1	1	1	1	1	1	1	1	8
progênie 6	1	1	1	1	1	1	1	1	8
progênie 7	1	1	1	1	1	1	1	1	8
progênie 8	1	1	1	1	1	1	1	1	8
progênie 9	1	1	1	1	1	1	1	1	8
progênie 10	1	1	1	1	1	1	1	1	8

Plano A

Genótipo	local 1	local 2	local 3	local 4	local 5	local 6	local 7	local 8	total
test 1	1	1	1	1	1	1	1	1	8
test 2	1	1	1	1	1	1	1	1	8
test 3	1	1	1	1	1	1	1	1	8
test 4	1	1	1	1	1	1	1	1	8
test 5	1	1	1	1	1	1	1	1	8
test 6	1	1	1	1	1	1	1	1	8
test 7	1	1	1	1	1	1	1	1	8
test 8	1	1	1	1	1	1	1	1	8
test 9	1	1	1	1	1	1	1	1	8
test 10	1	1	1	1	1	1	1	1	8
progênie 1	1		1			1		1	4
progênie 2	1	1		1			1		4
progênie 3		1	1		1			1	4
progênie 4	1		1			1		1	4
progênie 5					1	1	1	1	4
progênie 6	1	1		1			1		4
progênie 7			1	1		1	1		4
progênie 8		1		1	1	1			4
progênie 9	1	1			1			1	4
progênie 10			1	1	1		1		4

Plano B

Fonte: Do autor (2025)

Na Figura 1, observa-se que o plano A representa o cenário balanceado, em que as dez testemunhas e as dez progênes foram testadas nos oito locais. Por outro lado, no plano B é representado um cenário desbalanceado no qual as progênes sob teste estão presentes em apenas 50% dos locais, promovendo uma redução na mesma proporção no número de parcelas para testar estas progênes. Salienta-se que em ambos os planos ou cenários, as testemunhas estão presentes em todos os locais.

É importante destacar que para a geração dos cenários de desbalanceamento dos ensaios multiambientais, o modelo levou em consideração o fator família para otimizar o desbalanceamento e garantir que cada família fosse testada em todos os locais. Deste modo, considerando novamente a Figura 1 e admitindo uma família com 10 progênes testadas em oito locais com 50% de desbalanceamento, tem-se essa restrição atendida. A progênie 1 foi alocada nos locais 1, 3, 6 e 8, enquanto a progênie 2 foi testada nos locais 1, 2, 4 e 7 e, assim, sucessivamente. De modo que cada progênie fosse testada em apenas metade dos locais, mas que a família estivesse representada em todos os ambientes.

Os experimentos foram alocados em fazendas referências de cada região, áreas relativamente homogêneas e com adequada fertilidade de solo. O manejo do experimento foi de acordo com o manejo adotado pelo produtor rural, caracterizando uma situação real de produção agrícola.

3.2 Análise estatística dos dados

As análises estatísticas foram realizadas usando a abordagem de modelos mistos (HENDERSON, 1984). Os componentes de variância foram estimados pelo método de máxima verossimilhança residual (REML) (PATTERSON & THOMPSON 1971).

Os dados dos cenários de desbalanceamento dos ensaios multiambientais foram analisados de acordo com o seguinte modelo:

$$y = 1\mu + X_l\theta_l + Z_bu_b + Z_gu_g + Z_{gl}u_{gl} + e$$

em que: y é o vetor $n \times 1$ dos dados das n parcelas; 1 é o vetor $n \times 1$ de uns; μ é uma constante associada a todos os dados; θ_l é o vetor de efeitos fixos dos locais; u_b é o vetor de efeitos aleatórios dos blocos dentro dos locais, $u_b \sim N(0, I\sigma_b^2)$; u_g é o vetor de efeitos aleatórios dos genótipos, $u_g \sim N(0, G\sigma_g^2)$, sendo G a matriz de relacionamento genético.; u_{gl} é o vetor de efeitos aleatórios da interação genótipos por locais, $u_{gl} \sim N(0, I\sigma_{gl}^2)$; e é o vetor de efeitos

aleatórios dos erros, $e \sim N(0, I\sigma_e^2)$. X_l , Z_b , Z_g e Z_{gl} são as matrizes de delineamento associadas aos efeitos θ_l , u_b , u_g e u_{gl} , respectivamente. σ_b^2 , σ_g^2 , σ_{gl}^2 e σ_e^2 são os componentes de variância dos blocos dentro de repetições, genótipos, interação genótipos por ambientes e erros.

Foram obtidas as predições BLUP (*best linear unbiased prediction*) dos efeitos aleatórios e as estimativas BLUE (*best linear unbiased estimator*) dos efeitos fixos. Todos os cenários foram analisados considerando duas nuances de modelo: modelo ABLUP com a inclusão da matriz parentesco de Wright (A) baseada no pedigree ou genealogia, ou seja, a matriz G corresponde a matriz A; e o modelo IBLUP sem a inclusão de parentesco, ou seja, admitindo genótipos não aparentados, sendo a matriz G correspondente à matriz identidade da dimensão do número de genótipos. A partir do modelo ajustado, foram estimados os seguintes parâmetros:

- **Acurácia Seletiva** na média de progênie estimada a partir da seguinte expressão (RESENDE; ALVES, 2022):

$$r_{\hat{g}g} = \sqrt{1 - \frac{PEV}{\sigma_g^2}},$$

em que: PEV representa a variância média do erro de predição associada ao BLUP de progênie e σ_g^2 é a estimativa da variância genética (SEARLE; CASELLA; McCULLOCH, 1992).

- **Correlação Classificatória de Spearman** entre as predições BLUP dos cenários desbalanceado e balanceado.
- **Índice de Coincidência da Seleção**, proposto por Hamblin e Zimmermann (1986), utilizando a seguinte expressão:

$$IC = \left(\frac{A - C}{M - C} \right)$$

em que: A: número de progênies coincidentes nos dois cenários; M: número de progênies selecionadas em dado cenário, admitindo uma fração de selecionados de 10%; C: quantifica o número de progênies coincidentes devido ao acaso. Considerando a seleção das 10% melhores progênies, $C = A \times 10\%$.

As análises estatísticas dos dados foram realizadas usando o pacote ASReml – R 4.2 (GILMOUR et al. 2023) no software R (R Core Team 2023).

4 RESULTADOS E DISCUSSÃO

O número de progênies e testemunhas permaneceu constante em todos os cenários (Tabela 3), enquanto o desbalanceamento progressivo reduziu o número de parcelas de 4.320 (Cenário 0) para 540 (Cenário 7). Essa redução otimizou o plano experimental ao economizar tempo, mão de obra, custos logísticos e área experimental, viabilizando o experimento em espaços menores. Contudo, o desbalanceamento pode aumentar a variância residual, comprometendo a precisão das predições dos valores genéticos das progênies. Assim, é essencial equilibrar a minimização da perda de acurácia seletiva com os benefícios da otimização experimental, especialmente nas etapas iniciais de melhoramento, que envolvem a avaliação de muitos genótipos. Montesinos-López et al. (2023) destacaram as vantagens econômicas do desbalanceamento, que permitem testar o mesmo número de genótipos com menor custo ou aumentar o número de genótipos avaliados dentro do mesmo orçamento de um experimento balanceado. Esses resultados reforçam a viabilidade do desbalanceamento planejado como estratégia eficiente para otimizar recursos em programas de melhoramento genético (PERSA et al., 2023).

Tabela 3 - Detalhamento do número de progênies, testemunhas, locais e parcelas para cada cenário de desbalanceamento.

Itens	Cenário 0	Cenário 1	Cenário 2	Cenário 3	Cenário 4	Cenário 5	Cenário 6	Cenário 7
# Progênies	450	450	450	450	450	450	450	450
# Testemunhas	10	10	10	10	10	10	10	10
# Locais	8	8	8	8	8	8	8	8
Desbalanceamento (%)	0%	12,5%	25,0%	37,5%	50,0%	62,5%	75,0%	87,5%
# parcelas	4320	3780	3240	2700	2160	1620	1080	540

Fonte: Do autor (2025)

Ao analisar as estimativas dos parâmetros genéticos para os cenários estudados (Tabela 4), notou-se que a variância genética (σ_g^2) aumentou gradativamente à medida que o nível de desbalanceamento foi acentuado, tanto para o ABLUP quanto para o IBLUP. Conforme esperado, a variância do erro de predição (PEV) aumentou com o aumento do desbalanceamento em ambos os modelos com e sem a inclusão da informação de parentesco. Isso indica que a incerteza nas predições genéticas cresce em cenários com maior desbalanceamento. O aumento da PEV com o desbalanceamento é mais acentuado no IBLUP em comparação ao ABLUP, indicando que o modelo ABLUP apresenta maior robustez em

cenários com menor quantidade de dados genéticos. Cenários com PEV elevada (cenário 7, especialmente) refletem maior incerteza nas predições genéticas.

A acurácia seletiva (r_{gg}) diminuiu progressivamente em ambos os modelos (ABLUP e IBLUP) à medida que o desbalanceamento dos dados aumentou, refletindo em perda de precisão na seleção em cenários com maior grau de desbalanceamento. No método ABLUP, a acurácia variou de 0,67 a 0,79, enquanto no IBLUP, a variação foi de 0,53 a 0,75. De acordo com Resende e Duarte (2007), valores de acurácia superiores a 0,7 são considerados altos, enquanto valores entre 0,5 e 0,69 são classificados como moderados.

Tabela 4 - Estimativas de parâmetros genéticos e ambientais para produtividade referente à avaliação de progênies de soja sob diferentes cenários de desbalanceamento em ensaios multiambientais no delineamento em blocos aumentados.

Parâmetros	Cenários							
	0	1	2	3	4	5	6	7
ABLUP								
σ_g^2	0,0806	0,0848	0,0919	0,0917	0,1221	0,1191	0,1517	0,2136
PEV	0,0308	0,0335	0,0375	0,0393	0,0520	0,0555	0,0752	0,1186
r_{gg}	0,79	0,78	0,77	0,76	0,76	0,73	0,71	0,67
IBLUP								
σ_g^2	0,0792	0,0807	0,0881	0,0880	0,1080	0,1066	0,1334	0,1825
PEV	0,0342	0,0371	0,0422	0,0449	0,0570	0,0621	0,0840	0,1312
r_{gg}	0,75	0,74	0,72	0,70	0,69	0,65	0,61	0,53

σ_g^2 : variância genética de progênies, PEV: variância média do erro de predição associada ao BLUP de progênie; r_{gg} : acurácia seletiva na média de progênie.

Fonte: Do autor (2025).

Pelos resultados da Tabela 4 evidenciou-se que, embora ambos os métodos sejam afetados pelo desbalanceamento, o ABLUP manteve uma acurácia mais estável, enquanto o IBLUP apresentou uma redução mais acentuada, especialmente em situações de maior desbalanceamento. Jarquin et al. (2020) relataram resultados semelhantes em estudos com dados de milho, em que a redução no número de genótipos comuns avaliados em múltiplos ambientes levou a uma queda na acurácia preditiva.

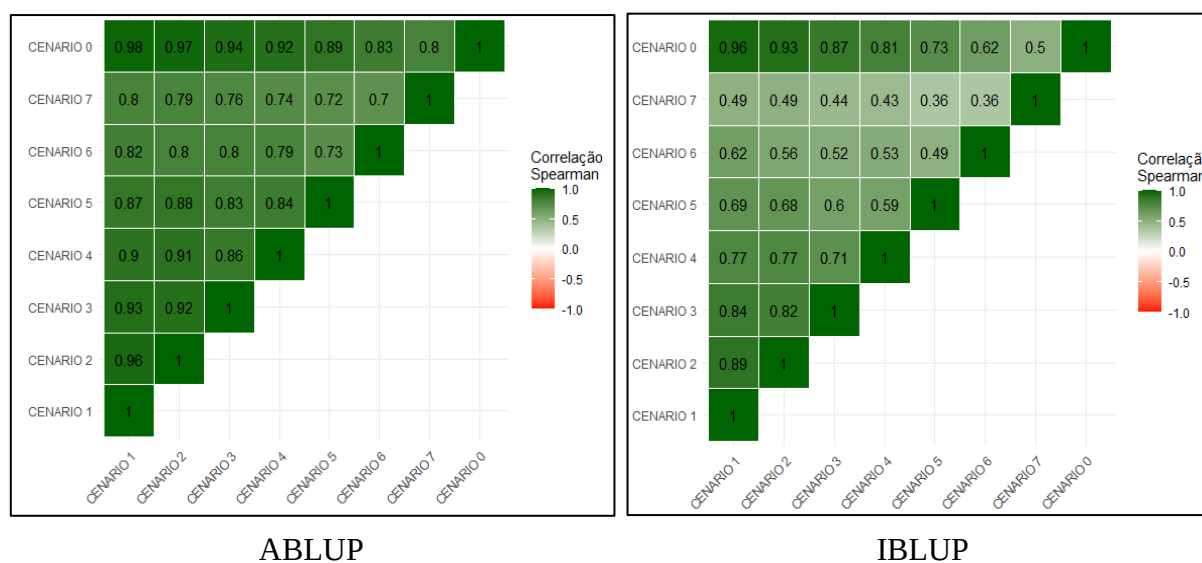
Este resultado também evidencia que a redução da acurácia seletiva foi relativamente baixa, mesmo em cenários mais desbalanceados. Estudos como o de Montesinos-López et al. (2023), que aplicaram métodos de desbalanceamento para conjuntos de dados de milho e trigo, demonstraram que mesmo em cenários extremos de desbalanceamento, a perda de acurácia preditiva foi baixa. Isso reforça que a otimização de planos experimentais em ensaios multiambientais por meio de planos experimentais desbalanceados é uma estratégia viável e

eficiente para avaliar muitos genótipos sem a necessidade de testá-los em todos os ambientes. Garcia-Abadillo et al. (2024), em estudos com cana-de-açúcar, e Jarquin et al. (2020), em milho, observaram que cenários com alta sobreposição de genótipos entre ambientes resultam apenas em ganhos marginais de acurácia preditiva. Esses autores sugerem que repetir apenas alguns genótipos em todos os ambientes, enquanto a maioria dos genótipos é avaliada de forma não sobreposta, é uma estratégia eficiente para maximizar o uso de recursos sem comprometer significativamente a precisão das predições. Essa abordagem é particularmente eficaz quando combinada com modelos que estimam a interação genótipo-ambiente (GxE) e incorporam dados de matriz de parentesco (PERSA et al., 2023; PERSA; IWATA; JARQUIN, 2020).

Segundo Jarquin et al. (2020), a integração de dados genômicos e de *pedigree* aumenta significativamente a capacidade preditiva dos modelos, mesmo em delineamentos desbalanceados. Essa estratégia tem sido amplamente aplicada no melhoramento de plantas autógamas, elevando a precisão das estimativas e facilitando a seleção de genótipos superiores (CROSSA et al., 2010; RESENDE et al., 2016; RUTKOSKI et al., 2016; VALENTE et al., 2016).

A influência do desbalanceamento na precisão do ranqueamento das progênes foi evidente nos resultados obtidos (Figura 2).

Figura 2 - Correlação classificatória de Spearman para produtividade de soja em oito cenários de desbalanceamento, comparando os métodos ABLUP e IBLUP.



Fonte: Do autor (2025).

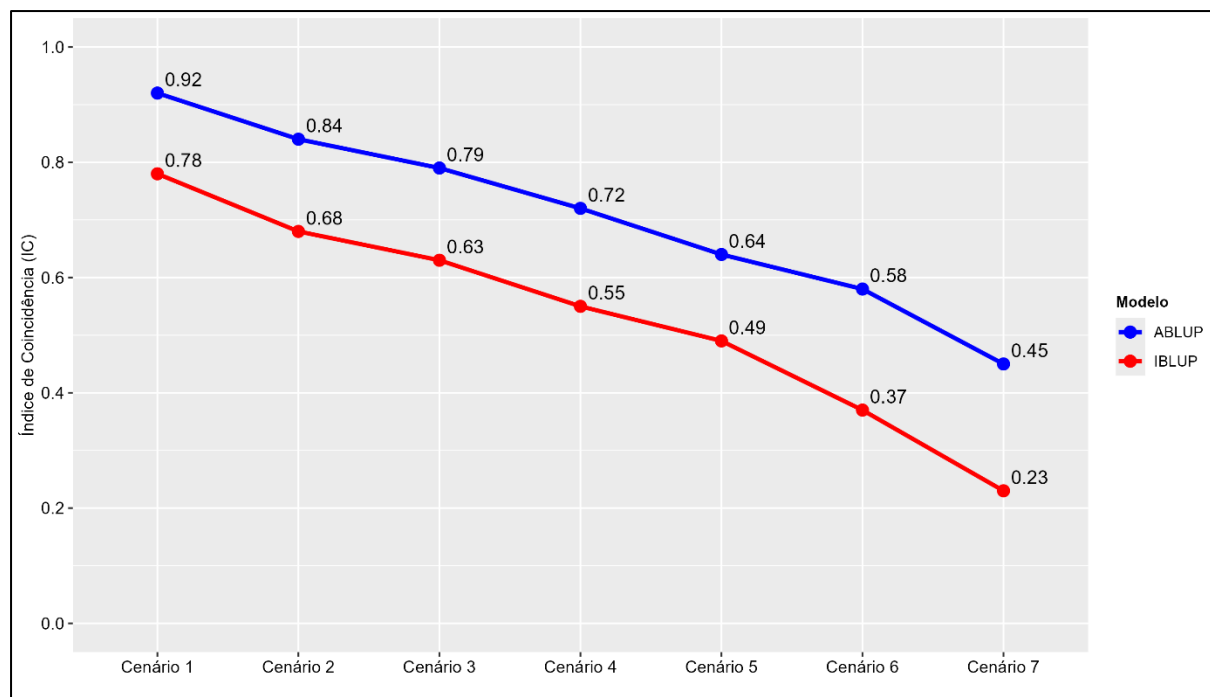
Com a metodologia ABLUP, observou-se uma leve redução na correlação de Spearman à medida que o desbalanceamento aumentou. No cenário 1, que apresenta menor desbalanceamento, a correlação foi alta (0,98), indicando que a ordenação dos genótipos permaneceu praticamente idêntica à do cenário balanceado. No entanto, essa correlação diminuiu gradualmente, atingindo 0,80 no cenário 7.

Para o método IBLUP, a redução na correlação de Spearman foi mais acentuada. No cenário 1, a correlação foi elevada (0,96), mas caiu para 0,50 no cenário 7, evidenciando uma grande mudança na ordenação dos genótipos entre o cenário balanceado e os desbalanceados mais extremos. Esses resultados reforçam que o REML/BLUP é uma metodologia eficaz para lidar com experimentos desbalanceados, garantindo uma ordenação estável dos genótipos. Contudo, quando o desbalanceamento é extremo e não há informações auxiliares de estrutura de família, a precisão da ordenação pode ser significativamente afetada.

Estes resultados corroboram com o estudo de Santos (2014), que avaliou o impacto do desbalanceamento na estimativa de parâmetros genéticos e observou que a correlação de Spearman entre valores genotípicos reais e preditos pelo método REML/BLUP permaneceu alta, mesmo com até 20% de desbalanceamento. De acordo com Carvalho, Lorencetti e Benin (2004), correlações entre 0,60 e 0,90 são classificadas como fortes, enquanto valores entre 0,30 e 0,60 são considerados moderados. Nesse contexto, os resultados indicam que, em cenários de desbalanceamento moderado, a correlação permaneceu forte, sugerindo que a ordenação dos genótipos ainda é confiável. No entanto, em cenários com desbalanceamento mais severo e sem o auxílio de informações de pedigree, a correlação caiu para níveis moderados (0,50), comprometendo a precisão do ranqueamento.

As estimativas do índice de coincidência (IC) decresceram com o aumento do nível desbalanceamento, independentemente do modelo de análise (Figura 3). No entanto, o ABLUP apresentou valores mais elevados de IC em todos os cenários, evidenciando maior capacidade de manter a seleção consistente em relação ao cenário balanceado.

Figura 3 - Índice de coincidência na seleção dos 10% melhores genótipos para sete cenários de desbalanceamento comparados ao cenário 0 (balanceado).



Fonte: Do autor (2025).

Nos cenários com menor desbalanceamento (cenário 1), o IC ainda é elevado, com valores de 92% e 78% para ABLUP e IBLUP, respectivamente, indicando que a seleção das 10% melhores progênies ainda é bastante semelhante ao cenário ideal (Figura 3). Contudo, à medida que o desbalanceamento aumenta, a consistência na seleção se reduz progressivamente. Isso é evidenciado pelos valores de IC mais baixos nos cenários intermediários, com valores de 72% e 55% no cenário 4, e ainda mais nos cenários altamente desbalanceados, com valores 45% e 23% no cenário 7, respectivamente, para os modelos ABLUP e IBLUP.

Os resultados destacam a importância de se considerar cuidadosamente o grau de desbalanceamento em experimentos conduzidos em vários ambientes. Portanto, o desbalanceamento deve ser utilizado como uma estratégia de otimização, mas com limites. É essencial equilibrar a redução de custos com a necessidade de manter uma precisão na seleção, para garantir que os ganhos genéticos não sejam comprometidos no longo prazo.

5 CONCLUSÕES

O desbalanceamento de progênies em ensaios multiambientes impacta na precisão das estimativas de parâmetros genéticos, na acurácia dos valores preditos das progênies e no ranqueamento das progênies para fins de seleção.

Planos experimentais com níveis moderados de desbalanceamento de progênies proporcionam uma otimização da alocação de recursos e podem ser indicados, especialmente, na fase inicial do programa de melhoramento, na qual há muitos genótipos sob teste e, comumente, limitação de material vegetal.

A definição do grau de desbalanceamento é um fator crítico para equilibrar a otimização de recursos e minimizar perdas na acurácia seletiva.

A obtenção de planos experimentais de ensaios multiambiente com desbalanceamento planejado levando em consideração a matriz de parentesco e sua posterior inclusão na análise contribui para a melhoria da precisão das estimativas dos parâmetros genéticos e a confiabilidade da seleção dos genótipos.

REFERÊNCIAS

- AKINWALE, R.O. et al. Effectiveness of different alpha lattice designs in the evaluation of maize (*Zea mays* L.) genotypes in a rainforest agro-ecology. **Heliyon**. 2021.
- ALLEN, T. et al. Soybean Yield Loss Estimates Due to Diseases in the United States and Ontario, Canada, from 2010 to 2014. **Plant Health Progress**, v. 18, n. 1, p. 19-27, 2017.
- ALMEIDA, L. A. de et al. Melhoramento da soja para regiões de baixas latitudes. In: QUEIROZ, M.A. de; GOEDERT, C.O.; RAMOS S.R.R. (Ed.). **Recursos genéticos e melhoramento de plantas para o Nordeste brasileiro**. Petrolina: Embrapa Semi-Árido; Brasília-DF: Embrapa Recursos Genéticos e Biotecnologia, nov. 1999.
- AMABILE, R. F.; VILELA, M. S.; PEIXOTO, J. R.(ed.) **Melhoramento de plantas: variabilidade genética, ferramentas e mercado**. Brasília, DF: Proimpress; Sociedade Brasileira de Melhoramento de Plantas, 2018. 108 p.
- ARAVIND J., MUKESH SANKAR S., WANKHEDE D.P., KAUR V. (2023). augmentedRCBD: Analysis of Augmented Randomised Complete Block Designs. R package version 0.1.7 <https://aravind-j.github.io/augmentedRCBD/> <https://cran.r-project.org/package=augmentedRCBD>.
- ATANDA, S.A. et al. Sparse testing using genomic prediction improves selection for breeding targets in elite spring wheat. **Theoretical and Applied Genetics**, v. 135, p. 1939–1950, 2022.
- BAILEY, R. A. **Design of Comparative Experiments**. Cambridge: Cambridge University Press, 2008.
- BIANCHI, M.C. et al. Heritability and the genotype $\times$ environment interaction in soybean. **Crop Science**, v. 3, 2020.
- BONATO, E. R.; BONATO, A.L. V. **A soja no Brasil: história e estatística**. Londrina, EMBRAPA-CNPSo, 1987, 61 p.
- BORÉM, A.; MIRANDA, G.V.; NETO FRITSCH, R. **Melhoramento de Plantas**. 8. ed. São Paulo: Editora Oficina do Texto, 2021. 384p.
- BUTLER, D. G.; SMITH, A. B.; CULLIS, B. R. On the Design of Field Experiments with Correlated Treatment Effects. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 19, n. 4, p. 541–557, 2014.
- CARPINETTI, L. C.R. **Planejamento e análise de experimentos**. Universidade de São Paulo, Escola de Engenharia de São Carlos, Departamento de Engenharia De Produção, São Carlos, 2009.
- CARVALHO, F.I.F.; LORENCETTI, C.; BENIN, G. **Estimativas e implicações da correlação no melhoramento vegetal** Pelotas: UFPel, 2004. 142p.

CARVALHO, M. P. et al. Adaptability and stability of conventional soybean by GGE biplot analysis. **Pesquisa Agropecuária Tropical**, Goiânia, v. 51, 2021.

CLEWER, A. G.; SCARISBRICK, D. H. **Practical Statistics and Experimental Design for Plant and Crop Science**. Chichester - UK: John Wiley & Sons Ltd, 2001. 346 p.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). Boletim da Safra de Grãos. Disponível em <<https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos>>. Acesso em 02/10/2023.

CRESPO-HERRERA L. et al. Genome-enabled prediction for sparse testing in multi-environmental wheat trials. **Plant Genome**, v. 14, n. 3, 2021

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.

CUCOLOTTO, M. et al. Genotype x environment interaction in soybean: evaluation through three methodologies. **Crop Breeding and Applied Biotechnology**, v. 7, p. 270-277, 2007.

CULLIS, B. R. et al. The Design of Early-Stage Plant Breeding Trials Using Genetic Relatedness. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 25, n. 4, p. 553–578, 2020.

CULLIS, B. R.; SMITH, A. B.; COOMBES, N. E. On the Design of Early Generation Variety Trials With Correlated Data. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 11, n. 4, p. 381–393, 2006.

DALL'AGNOL, A. A soja no brasil: evolução, causas, impactos e perspectivas. In: CONGRESO DE LA SOJA DEL MERCOSUR, 5.; FORO DE LA SOJA ASIA, 1., 2011, Rosário. Un grano: un universo. [Rosário: Asociación de la Cadena de la Soja Argentina], 2011. 4 p. 1 CD-ROM. MERCOSOJA, 2011.

DONG, L et al. Genetic basis and adaptation trajectory of soybean from its temperate origin to tropics. **Nat Commun**, v. 12, 2021.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). Cultivos de Soja. Disponível em <<https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/soja/pre-producao/caracteristicas-da-especie-e-relacoes-com-o-ambiente/cultivares>>. Acesso em 01/10/2023

FEDERER, W.T. Augmented (hoonuiaku) designs. **Hawaiian Planters' Record**, Aiea, v.55, p.191-208, 1956

FREITAS, A.R. **Estatística experimental na agropecuária**. 1. ed. Brasília, DF: Embrapa, 2022.

GARCIA-ABADILLO, J. et al. Sparse testing designs for optimizing predictive ability in sugarcane populations. **Frontiers in Plant Science**, v. 15, 2024.

GAUCH, H.G. A Simple Protocol for AMMI Analysis of Yield Trials. **Crop Science**, v. 53, n. 5, p. 1860-1869, 2013.

GESTEIRA, G. de S. et al. Selection of Early Soybean Inbred Lines Using Multiple Indices. **Crop Science**, v. 58, 2018.

GILMOUR, A. R., Gogel, B. J., Cullis, B. R., Thompson, R. (2023). ASReml User Guide Release 4.1 Structural Specification. **VSN International Ltd.**, Hemel Hempstead, HP1 1ES, UK.

HAMBLIN, J.; ZIMMERMANN, M. J. de O. Breeding Common Bean for Yield in Mixtures. In: JANICK, **Plant Breeding Reviews**. p. 245-272. 1986.

HENDERSON, C. R. A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. **Biometrics**, v. 32, n. 1, 1976.

HENDERSON, C.R. **Applications of linear models in animal breeding**. Guelph: University of Guelph, 1984. 462p.

HOEFLER, R. et al. Do Spatial Designs Outperform Classic Experimental Designs?. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 25, n. 4, p. 523–552, 2020.

JARQUÍN, D. et al. (2020). Genomic prediction enhanced sparse testing for multi-environment trials. **G3 - Genes, Genomes, Genetics**, v. 10, n. 8, p. 2725–2739, 2020.

KASTER, M.; FARIAS, J.R.B. Regionalização dos testes de Valor de Cultivo e Uso e da indicação de cultivares de soja - Terceira aproximação. Londrina: Embrapa Soja, 2012. 69 p. ISSN 2176-2937.

LAWSON, J. **Design and Analysis of Experiments with R**. Provo, Utah – USA: Chapman and Hall/CRC, 2015. 620 p.

LI M. et al. Comprehensive mapping of abiotic stress inputs into the soybean circadian clock. **Proceedings of the National Academy of Sciences**, v. 116, n. 47, 2019.

LIN, F., et al. Breeding for disease resistance in soybean: a global perspective. **Theoretical and Applied Genetics**, 2022.

MARANNA, S. et al. Breeding for higher yield, early maturity, wider adaptability and waterlogging tolerance in soybean (*Glycine max* L.): A case study. **Scientific Reports**, v. 11, 2021.

- MARTIN, T. N. et al. Definição de plano experimental para comparação de cultivares de milho em áreas limitadas. **Revista Brasileira de Milho e Sorgo**, v.3, n.3, p.325-332, 2004.
- MARTIN, T. N. et al. Metodologia experimental para rendimento de grãos de soja em condições de restrição de espaço. **Bragantia**, Campinas, v. 66, n. 3, p.521-526, 2007
- MOEHRING J.; WILLIAMS E.R.; PIEPHO H.P. Efficiency of augmented p-rep designs in multi-environmental trials. **Theoretical and Applied Genetics**, v. 127, p. 1049-1060, 2014.
- MONA, I. et al. Using of alpha lattice design for increasing precision of faba bean yield trials. **Journal of Environmental Science**, v. 44, n. 2, 2018.
- MONTESINOS-LÓPEZ, O.A. et al. Optimizing Sparse Testing for Genomic Prediction of Plant Breeding Crops. **Genes**, v. 14, n. 4, 2023.
- MURPHY, S. E. et al. Genotype × Environment Interaction and Stability for Isoflavone Content in Soybean. **Crop Science**, v. 49, p. 1313-1321, 2009.
- MUTAVA, R.N. et al. Understanding abiotic stress tolerance mechanisms in soybean: A comparative evaluation of soybean response to drought and flooding stress. **Plant Physiology and Biochemistry**, v. 86, 2015, p. 109-120.
- PACHECO, R. M. et al. Environmental stratifications for soybean cultivar recommendation and its consistency over time. **Genetics and Molecular Research**, vol. 16, n.3, 2017.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v.58, p.545-554, 1971.
- PEREIRA, D.R. et al. Genotype x environment interactions and implications for associations among soybean traits. **Genetics and Molecular Research**, vol. 18, n.3, 2019.
- PERSA, R. et al. Improving predictive ability in sparse testing designs in soybean populations. **Frontiers in Plant Science**, v. 14, 2023.
- PERSA, R.; IWATA, H.; JARQUIN, D. Use of family structure information in interaction with environments for leveraging genomic prediction models. **The Crop Journal**, v. 8, p.843-854, 2020.
- PETERNELLI, L.A. et al. Delineamentos aumentados no melhoramento de plantas em condições de restrições de recursos. **Ciência Rural**, v. 39, n. 9, p. 2425–2430, 2009.
- R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- RAMALHO, M. A. P. et al. **Genética na Agropecuária**. 5. ed. Lavras: Editora UFLA, 2012, 566 p.

RAMALHO, M. A. P.; FERREIRA, D. F.; DE OLIVEIRA, A. C. **Experimentação em Genética e Melhoramento de Plantas**. 3. ed. Lavras: Editora UFLA, 2012, 305 p.

RAMALHO, M. A. P.; MARQUES, T. L.; LEMOS, R. do C. Plant breeding in Brazil: Retrospective of the past 50 years. **Crop Breeding and Applied Biotechnology**, v. 21, 2021.

RAMOS JUNIOR, E.U. et al. Identification of soybean genotypes with high stability for the Brazilian macro-region 402 via biplot analysis. **Genetics and Molecular Research**, vol. 1, n.3, 2017.

RANI, R. et al. Analysis of genotype $\times$ environment interactions for agronomic traits of soybean (*Glycine max* [L.] Merr.) using association mapping. **Frontiers in Genetics**. v. 13, 2023.

RESENDE, M. D. V.; ALVES, R. S. Statistical significance, selection accuracy, and experimental precision in plant breeding. **Crop Breeding and Applied Biotechnology**, v. 22, n. 3, 2022.

RESENDE, M. D. V. de, et al. Selection Index with Parents, Populations, Progenies, and Generations Effects in Autogamous Plant Breeding. **Crop Science**, v. 56, p. 530-546, 2016.

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

RUPESH, D. et al. Integrating omic approaches for abiotic stress tolerance in soybean. **Frontiers in Plant Science**. v. 5, 2014.

RUTKOSK, J. et al. Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. **G3: Genes, Genomes, Genetics**, v. 6, 2016.

SANTOS, C.F. dos. **Uso de diferentes metodologias estatísticas no melhoramento do amendoim forrageiro**. 2014. Dissertação (Mestrado em Produção Vegetal) – Universidade Federal do Acre, Rio Branco, 2014.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. **Variance components**. New York: John Wiley & Sons, 1992. 501p.

SILVA, R.R.; BENIN, G. Análises Biplot: conceitos, interpretações e aplicações. **Ciência Rural**, v. 42, n. 8, p. 1404-1412, 2012.

SOARES, I.O. et al. Stability and adaptability of soybean cultivars in Minas Gerais. **Genetics and Molecular Research**, vol. 16, n.3, 2017.

SOUZA, R.R. de. et al. Soybean grain yield in highland and lowland cultivation systems: A genotype by environment interaction approach. **Annals of Applied Biology**, v.179, p.302-318, 2021.

SQUILASSI, M.G. **Interação de genótipos com ambientes**. 1. ed. Aracajú: Embrapa Tabuleiros Costeiros, 2003, 47p. Disponível em: <http://www.cpatc.embrapa.br>.

STURION, J.A.; RESENDE, M. D. V.; Eficiência do Delineamento Experimental e Capacidade de Teste no Melhoramento Genético da Erva-Mate (*Ilex paraguariensis* St. Hil.). **Pesquisa Florestal Brasileira**. Colombo, n. 50, p. 3-10, 2005.

UNITED STATES DEPARTMENT OF AGRICULTURE (USDA). World Agricultural Production. Disponível em <<https://apps.fas.usda.gov/psdonline/circulars/production.pdf>>. Acesso em 02/10/2023.
VALENTE, M. S. F. et al. Seleção genômica para melhoramento vegetal com diferentes estruturas populacionais. **Pesq. agropec. bras.**, v. 51, n. 11, 2016.

VIANA, J. S. et al. Precocidade de cultivares de soja em sistemas de cultivo com milho em Areia, Estado da Paraíba. **Acta Scientiarum. Agronomy**. Maringá, v. 31, n. 3, p. 481-488, 2009.

YAN, W. Analysis and Handling of $G \times E$ in a Practical Breeding Program. **Crop Science**, v. 56, 2016.

YAN, W. et al. GGE Biplot vs. AMMI Analysis of Genotype-by-Environment Data. **Crop Science**, v. 47, p. 643-653, 2007.

APÊNDICE A - PRODUTO TÉCNICO ATRELADO

Este guia explica como criar um delineamento experimental em que as linhagens são distribuídas em um número específico de locais, com um controle preciso do desbalanceamento. O script foi desenvolvido em **R** e pode ser adaptado para diferentes cenários de experimentação.

Passo 1: Definir os Parâmetros do Experimento

Antes de começar, defina os parâmetros básicos do seu experimento:

- **Número de linhagens:** Quantas linhagens você deseja testar.
- **Número de locais:** Quantos locais estão disponíveis para o experimento.
- **Número de checks:** Quantas testemunhas (*checks*) serão usadas em cada bloco.
- **Percentual de desbalanceamento:** A proporção de locais em que cada linhagem será testada (por exemplo, 50% significa que cada linhagem aparecerá em metade dos locais).

Exemplo:

```
#Rotina_R  
num_linhagens <- 200 # 200 linhagens  
num_locais <- 10    # 10 locais  
linhagens_por_bloco <- 50 # (Delineamento em Blocos Aumentados)  
num_checks <- 10    # 10 checks por bloco  
percentual_desbalanceamento <- 0.5 # 50% de desbalanceamento sobre as linhagens
```

Passo 2: Calcular o Número de Locais por Linhagem

Com base no percentual de desbalanceamento, calcule quantos locais cada linhagem deve ocupar.

Exemplo:

```
locais_por_linhagem <- round(num_locais * percentual_desbalanceamento)
```

Para 10 locais e 50% de desbalanceamento, cada linhagem aparecerá em **5 locais**.

Passo 3: Criar a Estrutura do Delineamento

Crie um *dataframe* para armazenar a presença das linhagens em cada local.

Exemplo:

```
delineamento <- data.frame(  
  genotipo = paste0("linhagem_", 1:num_linhagens)  
)  
delineamento[, paste0("Local_", 1:num_locais)] <- 0
```

Passo 4: Distribuir as Linhagens nos Locais

Distribua aleatoriamente as linhagens nos locais, garantindo que cada linhagem apareça no número correto de locais.

Exemplo:

```
set.seed(123) # Para reprodutibilidade  
for (i in 1:num_linhagens) {  
  locais_sorteados <- sample(1:num_locais, locais_por_linhagem)  
  delineamento[i, locais_sorteados + 1] <- 1  
}
```

Passo 5: Verificar e Ajustar o Número de Linhagens por Local

Garanta que nenhum local tenha mais linhagens do que o permitido. O número máximo de linhagens por local é calculado como:

```
max_linhagens_por_local <- (num_linhagens * locais_por_linhagem) / num_locais
```

Se um local tiver mais linhagens do que o permitido, remova linhagens excedentes.

Exemplo:


```

for (local in 1:num_locais) {
  while (sum(delineamento[, local + 1]) > max_linhagens_por_local) {
    linhagem_para_remover <- which(delineamento[, local + 1] == 1 &
rowSums(delineamento[, 2:(num_locais + 1)]) > locais_por_linhagem - 1)
    if (length(linhagem_para_remover) > 0) {
      delineamento[linhagem_para_remover[1], local + 1] <- 0
    } else {
      stop("Não há linhagens disponíveis para remover do local ", local)
    }
  }
}
}

```

Passo 6: Subdividir Cada Local em Blocos

Divida cada local em blocos, onde cada bloco contém um número fixo de linhagens.

Exemplo:

```

library(tidyr)
library(dplyr)
delineamento_long <- delineamento %>%
  pivot_longer(cols = starts_with("Local_"), names_to = "Local", values_to = "Presente")
  %>%
  filter(Presente == 1) %>%
  select(-Presente)
delineamento_blocos <- delineamento_long %>%
  group_by(Local) %>%
  mutate(
    Bloco = rep(1:ceiling(n() / linhagens_por_bloco), each = linhagens_por_bloco, length.out
= n())
  ) %>%
  ungroup()

```

Passo 7: Adicionar *Checks* em Cada Bloco

Adicione as testemunhas (*checks*) no início de cada bloco.

Exemplo:

```
checks <- paste0("check_", 1:num_checks)
checks_df <- expand.grid(
  genotipo = checks,
  Local = paste0("Local_", 1:num_locais),
  Bloco = 1:ceiling(max(delineamento_blocos$Bloco))
)
delineamento_final <- delineamento_blocos %>%
  bind_rows(checks_df) %>%
  arrange(Local, Bloco, genotipo) %>%
  group_by(Local, Bloco) %>%
  mutate(Ordem = row_number()) %>%
  ungroup()
```

Passo 8: Preencher Parcelas Faltantes com *Checks*

Se o último bloco de um local não tiver 60 tratamentos (10 *checks* + 50 linhagens), preencha os espaços faltantes com *checks*.

Exemplo:

```
faltando <- delineamento_final %>%
  group_by(Local, Bloco) %>%
  summarise(n = n(), .groups = "drop") %>%
  filter(n < (num_checks + linhagens_por_bloco))
if (nrow(faltando) > 0) {
  for (i in 1:nrow(faltando)) {
    local <- faltando$Local[i]
    bloco <- faltando$Bloco[i]
    n_faltando <- (num_checks + linhagens_por_bloco) - faltando$n[i]
    checks_adicionais <- data.frame(
      genotipo = paste0("check_", sample(1:num_checks, n_faltando, replace = TRUE)),
      Local = local,
      Bloco = bloco,
      Ordem = NA
    )
    delineamento_final <- bind_rows(delineamento_final, checks_adicionais)
  }
  delineamento_final <- delineamento_final %>%
    arrange(Local, Bloco, Ordem) %>%
    group_by(Local, Bloco) %>%
    mutate(Ordem = row_number()) %>%
    ungroup()
}
```

Passo 9: Verificar e Salvar o Delineamento

Verifique se o delineamento está correto e salve os resultados em um arquivo CSV.

Exemplo:

```
parcelas_por_local_final <- delineamento_final %>%  
  group_by(Local) %>%  
  summarise(n = n(), .groups = "drop")  
  
print("Número de parcelas por local após o ajuste:")  
print(parcelas_por_local_final)  
  
write.csv(delineamento_final, "delineamento_final_com_checks.csv", row.names =  
FALSE)
```

RESULTADO FINAL

- Cada linhagem é testada em x locais de acordo com a taxa de desbalanceamento.
- Cada local tem blocos completos com **60 tratamentos** (10 checks + 50 linhagens).
- A redução no número de parcelas varia de acordo com o grau de desbalanceamento definido.
- O número de linhagens a ser testada deve estar em concordância com os parâmetros pré-estabelecidos para um bom desbalanceamento no delineamento experimental.