

JOÃO MARCOS AMARIO DE SOUSA

**MODELAGEM DA (CO)VARIÂNCIA GENÉTICA E RESIDUAL NA ANÁLISE DE
ENSAIOS MULTIAMBIENTES DE POPULAÇÕES SEGREGANTES DE TRIGO
TROPICAL**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

Orientador: Maicon Nardino

Coorientador: Kaio Olimpio das Graças Dias

**VIÇOSA - MINAS GERAIS
2024**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

S725m
2024

Sousa, João Marcos Amario de, 1999-
Modelagem da (co)variância genética e residual na análise de ensaios multiambientes de populações segregantes de trigo tropical / João Marcos Amario de Sousa. – Viçosa, MG, 2024.
1 dissertação eletrônica (30 f.): il. (algumas color.).

Orientador: Maicon Nardino.

Dissertação (mestrado) - Universidade Federal de Viçosa, Departamento de Agronomia, 2024.

Referências bibliográficas: f. 27-30.

DOI: <https://doi.org/10.47328/ufvbbt.2024.396>

Modo de acesso: World Wide Web.

1. Trigo - Melhoramento genético. 2. Interação genótipo-ambiente. 3. Modelos multiníveis (Estatísticas). 4. Análise de variância. I. Nardino, Maicon, 1988-. II. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Genética e Melhoramento. III. Título.

CDD 22. ed. 633.112

Bibliotecário(a) responsável: Bruna Silva CRB-6/2552


JOÃO MARCOS AMARIO DE SOUSA

**MODELAGEM DA (CO)VARIÂNCIA GENÉTICA E RESIDUAL NA ANÁLISE DE
ENSAIOS MULTIAMBIENTES DE POPULAÇÕES SEGREGANTES DE TRIGO
TROPICAL**


Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento para obtenção do título de *Magister Scientiae*.

APROVADA: 29 de fevereiro de 2024.

Assentimento:

Documento assinado digitalmente
 JOAO MARCOS AMARIO DE SOUSA
Data: 26/07/2024 11:20:38-0300
Verifique em <https://validar.iti.gov.br>

João Marcos Amario de Sousa
Autor

Documento assinado digitalmente
 MAICON NARDINO
Data: 26/07/2024 11:04:15-0300
Verifique em <https://validar.iti.gov.br>

Maicon Nardino
Orientador

A minha mãe Francisca Artur Amario.

AGRADECIMENTOS

A Deus, por sempre me iluminar nos momentos de aflição.

A minha mãe, Francisca Artur Amario, por sempre me apoiar.

A fazenda Mundo Novo (Nelore GGBF), do meu amigo Gilvan Gomes Barros Filho, pelo apoio a minha jornada de pós-graduação.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

Aos colegas do Programa Trigo UFV pela amizade.

Ao meu professor e orientador, Dr. Maicon Nardino, pelos ensinamentos e pelo exemplo de profissionalismo. Sempre instigando a buscar a excelência como cientista.

Ao meu coorientador, Dr. Kaio Olímpio das Graças Dias, pela disponibilidade em sempre ajudar e pelos ensinamentos.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

A Cooperativa Agropecuária do Alto Paranaíba (COOPADAP), São Gotardo, MG, pela condução dos ensaios.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

“Se eu vi mais longe, foi por estar sobre ombros de gigantes.”

(Sir Isaac Newton)

RESUMO

SOUSA, João Marcos Amario de, M.Sc., Universidade Federal de Viçosa, fevereiro de 2024. **Modelagem da (co)variância genética e residual na análise de ensaios multiambientais de populações segregantes de trigo tropical**. Orientador: Maicon Nardino. Coorientadores: Kaio Olimpio Das Graças Dias.

A quantificação da interação $G \times A$ pelo método dos quadrados mínimos ordinários considera as variâncias residuais ao longo dos ambientes homogêneos, o que pode não ocorrer na maioria dos casos, levando a estimativas viesadas dos componentes de variância. A análise de ensaios multiambientais via modelos mistos permite lidar com desbalanceamento genético e estatístico, e modelar as variâncias-covariâncias residuais ao longo dos ambientes, resultando em estimativas adequadas dos componentes de variância e predição acurada dos valores genotípicos. Com isso, o objetivo do trabalho foi implementar uma estrutura de modelagem das variâncias genéticas e residuais ao longo dos ambientes para uma análise de modelos mistos. Utilizamos dados de produtividade de grãos altura de planta e ciclo de três gerações em diferentes ambientes, conduzidos em condições de sequeiro e irrigado, em duas localidades no Estado de Minas Gerais. Os dados foram submetidos a análise individual para estimação dos componentes de variância via REML. Foi realizada a análise conjunta com base em diferentes modelos de modelagem dos efeitos genéticos e residuais para a predição dos valores genotípicos via REML/BLUP. Cada modelo teve seu ajuste testado pelo critério de informação de Akaike (AIC) e de Schwarz ou bayesiano (BIC). O modelo mais adequado para estimar os componentes de variância e os valores genéticos foi a estrutura de simetria composta heterogênea (CSH) para o efeito genotípico aliado ao de simetria diagonal (D) para o efeito residual.

Palavras-chave: Interação Genótipos por Ambientes. Modelos Mistos. Variâncias-Covariâncias Residuais e Genotípicas.

ABSTRACT

SOUSA, João Marcos Amario de, M.Sc., Universidade Federal de Viçosa, February, 2024. **Modeling genetic and residual (co)variance in the analysis of multi-environment trials of segregating populations of tropical wheat.** Adviser: Maicon Nardino. Co-advisers: Kaio Olimpio Das Graças Dias.

Analyzing G x E interaction by the ordinary least squares method assumes that residual variances across environments are homogeneous, which is often not the case, leading to biased estimates of variance components. Mixed models enable the handling of genetic and statistical imbalances in multi-environment trials (MET) analyses and allow for modeling residual (co)variances across environments, resulting in appropriate estimates of variance components and accurate prediction of genotypic values - BLUP. Therefore, the objective of this study was to implement a modeling structure for genetic and residual variances across environments for a mixed model analysis. We utilized data on grain yield (PROD), plant height (ALT), and days to heading (ESP) from three generations evaluated in different environments, conducted under rainfed and irrigated conditions, in two locations in the state of Minas Gerais, Brazil. The data underwent individual analysis to estimate variance components via REML. A joint analysis was conducted based on different models for modeling genetic and residual effects to predict genotypic values via REML/BLUP. Each model was assessed using the Akaike information criterion (AIC) and Schwarz or Bayesian information criterion (BIC). The most suitable model for estimating variance components and genetic values was the heterogeneous compound symmetry (CSH) structure for the genotypic effect combined with the diagonal (D) symmetry structure for the residual effect.

Keywords: Genotype by Environment Interaction. Mixed Models. Residual and Genotypic Variance-Covariances.

SUMÁRIO

1. INTRODUÇÃO	9
2. REVISÃO DE LITERATURA	10
2.1 Origem e dispersão do trigo	10
2.2 Importância socioeconômica	11
2.3 Melhoramento genético do trigo	12
2.4 Uso de Modelos Mistos	13
3. MATERIAL E MÉTODOS	13
3.1 Material genético e condução experimental	14
3.2 Análises estatísticas	15
3.2.1 Modelo para as análises individuais	15
3.2.2 Modelo para a análise conjunta	15
3.2.3 Teste de significância dos efeitos aleatórios	16
3.2.4 Parâmetros genéticos	16
3.2.5 Seleção de modelo	17
3.2.6 Coeficiente de coincidência	18
4. RESULTADOS E DISCUSSÃO	19
7. CONCLUSÕES	26
8. REFERÊNCIAS	27

1. INTRODUÇÃO

O trigo (*Triticum aestivum* L.) é um cereal com destacada importância e cultivado em todo o mundo (ALI et al., 2022), fornecendo 20% das calorias consumidas pela população mundial (SINGH et al., 2021) e sendo cultivado em 222,687 milhões de hectares (USDA, 2024). O Brasil é o 16º maior produtor mundial, com 3,45 milhões de hectares (USDA, 2024; CONAB, 2024). Entretanto, a produção nacional ainda não atende a demanda interna do mercado, mesmo diante do incremento na produtividade na ordem de 1.963 kg ha⁻¹, comparando com a década de 1940, quando começaram a ser lançados os primeiros materiais melhorados (LIMA et al., 2022). Com isso, de acordo com a estimativa da Conab (2024), safra 23/24, o Brasil necessita importar mais da metade do trigo que é consumido no país.

Diante da crescente demanda de consumo e da necessidade de adaptação às recentes tendências de melhoramento (LEE et al., 2023), a investigação da interação genótipos por ambientes (GxA), estabilidade e adaptabilidade são essenciais na recomendação de genótipos superiores (PENG et al., 1999; HALL et al., 2013; LI et al., 2019). Portanto, com o desenvolvimento de cultivares mais resilientes e produtivas, promovem avanços cruciais na busca da autossuficiência na produção de trigo brasileiro. Para Júnior et al. (2024), o Brasil pode alcançar a autossuficiência em trigo, apesar das mudanças climáticas, adotando estratégias de combinação tecnológicas inovadoras.

A interação GxA é essencial para o progresso de todo programa de melhoramento (VAN EEUWIJK et al., 2016) e representa um desafio para os melhoristas (UCHENDU et al., 2022), mas permite a compreensão da base genética de adaptação frente a variações ambientais (HUDSON et al., 2022). Compreender a importância relativa da GEI otimiza o programa de melhoramento (ROGERS et al., 2021). Essas interações representam uma parte da variância fenotípica que impede a adaptação ampla, mas podem ser exploradas para adaptações específicas (FALCONER e MACKAY, 1996), tornando possível identificar e recomendar cultivares estáveis e adaptadas às diversas condições ambientais (EGEA-GILBERT et al., 2021).

Normalmente, a GxA é avaliada ao final dos programas de melhoramento, para recomendar genótipos com adaptação específica. No entanto, vários fatores abióticos e bióticos influenciam a expressão de genes que controlam as

características agronômicas (UCHENDU et al., 2022), e avaliar materiais nos estágios finais do programa pode ser um desafio. Estimar GxA em populações permite ao melhorista verificar o desempenho agronômico ao longo das gerações, fornecendo informações mais precisas sobre as estimativas dos componentes de variância, auxiliando na identificação das populações mais promissoras. Prever como os genótipos respondem às mudanças no ambiente seletivo é crucial para os programas de melhoramento (HÄMÄLÄ et al., 2022).

Em populações segregantes, pode haver desequilíbrio genético ao longo das gerações, o que dificulta a análise conjunta por abordagens estatísticas comumente usadas para estudar a GxA (SMITH et al., 2005; GEZAN et al., 2017). Mesmo em linhagens ou híbridos simples, quando há uma GxA significativa, há expressão diferencial dos genes, o que se reflete no valor fenotípico (REIS et al., 2020; UCHENDU et al., 2022). Assumir a pressuposição da análise via mínimos quadrados ordinários, onde as variâncias genéticas e residuais são homogêneas, pode levar a estimativas tendenciosas em uma análise de variância.

Métodos clássicos não consideram as covariâncias dos efeitos genéticos e não genéticos dos ambientes (DIAS et al., 2020), mesmo diante de dados desbalanceados genética e estatisticamente. A análise de ensaios multiambientes (MET) via modelos mistos permite lidar com desbalanceamento genético e estatístico, modelando as (co)variâncias residuais e genotípicas ao longo dos ambientes, resultando em estimativas mais apropriadas. Uma análise via modelos mistos para levar em consideração a heterogeneidade de variâncias e covariâncias consideraria matrizes de variância-covariância não estruturadas para os efeitos genéticos e residuais em diferentes ambientes (GEZAN et al., 2017; MALOSETTI et al., 2013).

Com isso, o objetivo do trabalho foi comparar estruturas de modelagem das (co)variâncias residuais e genéticas ao longo dos ambientes para uma análise de modelos mistos, utilizando a metodologia Máxima Verossimilhança residual para estimar os componentes de variância e REML/BLUP.

2. REVISÃO DE LITERATURA

2.1 Origem e dispersão do trigo

O trigo (*Triticum aestivum* L., $2n=6x=42$) é uma espécie hexaplóide pertencente a tribo *Triticeae*, da família *Poaceae*, que se originou a partir da hibridação independente seguida de duplicações espontâneas dos cromossomos. A espécie *T. aestivum* (AABBDD) contém dois genomas homólogos aos genomas A e B da espécie *Triticum turgidum*. Huang et al. (2002) constataram que sequências do genoma D do *T. aestivum* e da espécie silvestre *Aegilops tauschii* são idênticas. Essa constatação sustenta a hipótese que *T. aestivum* é resultado da hibridação entre *T. turgidum* e *Ae. tauschii*, há cerca de 8 mil anos (PIANA e CARVALHO, 2008).

Os primeiros vestígios arqueológicos do trigo hexaplóide *T. aestivum* foram encontrados em Can Hasan, na Turquia, há 7 mil a.C. Entre 6 mil e 5 mil a.C., ganhou importância na agricultura irrigada das planícies da Mesopotâmia e oeste do Irã, expandindo-se para as bacias do rio Nilo e alcançando a Europa Central e Ocidental no final do quarto milênio a.C. (FELDMAN et al., 1995).

No continente americano, o trigo chegou após sua descoberta no século XV. Na América do Sul, foi plantada inicialmente na atual província de Santa Fé na Argentina em 1527. Indícios apontam que o trigo foi introduzido no Brasil por Martim Afonso da Silva em 1534, se espalhando pela capitania de São Vicente, acompanhando o deslocamento de famílias europeias (PIANA e CARVALHO, 2008; SOUSA e ROSA, 1985).

2.2 Importância socioeconômica

Atualmente, o trigo ocupa a posição de segundo cereal mais produzido no mundo, ficando atrás somente do milho. Com quase 786 milhões de toneladas em 2023, mantendo-se como uma das culturas com a maior área cultivada, abrangendo mais de 222 milhões de hectares. Os principais produtores mundiais são a China (com 136,5 milhões de toneladas), União Europeia (134 milhões de toneladas), Índia (110,5 milhões de toneladas), Rússia (91 milhões de toneladas) e Estados Unidos (49,3 milhões de toneladas). O Brasil, em 16º lugar no rank produz cerca de 8,1 milhões de toneladas (USDA, 2024). A segurança alimentar global está intrinsecamente ligada à triticultura, uma vez que desempenha um papel essencial na dieta humana, fornecendo 20% das calorias consumidas pela população mundial (SINGH et al., 2021).

O trigo foi introduzido no Brasil no século XVI, inicialmente destacando-se como exportador e tornando-se pioneiro (BORÉM e SCHEEREN, 2015). Contudo, atualmente, o país figura entre os principais importadores, já que a produção nacional ainda não atende à demanda interna, resultando na importação de mais da metade do trigo consumido no país (CONAB, 2024; USDA, 2024). Em 2023, aproximadamente 5,60 milhões toneladas de trigo foram importadas, colocando o Brasil em 10º lugar no ranking mundial de importadores (USDA, 2024).

2.3 Melhoramento genético do trigo

No geral, um programa de melhoramento de trigo passa por diversas etapas para a geração de material melhorado. Inicialmente, define-se o objetivo do programa. A segunda fase envolve a identificação de germoplasma de interesse e a introdução de variabilidade genética por meio da hibridação ou da introdução de germoplasma. A terceira etapa abrange a seleção e condução de populações para obtenção de linhagens, utilizando métodos apropriados. A quarta etapa consiste na avaliação das linhagens em diferentes ambientes e anos, aplicando todos os princípios da experimentação, para determinar seu comportamento. Por fim, a quinta etapa culmina no lançamento da cultivar desenvolvida.

O aumento da produtividade figura como uma prioridade central nos programas de melhoramento genético do trigo. Diversos projetos, liderados tanto por entidades públicas quanto privadas, têm se empenhado nesse propósito. Contudo, no contexto brasileiro, há uma atenção especial voltada para o desenvolvimento de genótipos que apresentem resistência às principais doenças: ferrugem (*Puccinia triticina*) e giberela (*Gibberella zeae*) na região sul, mancha amarela (*Pyrenophora tritici-repentins*) e brusone (*Magnaporthe oryzae pat. triticum*) na região central.

Outros atributos de destaque incluem ciclo curto, ampla adaptabilidade, estabilidade na produção de grãos e aptidão tecnológica para a indústria de panificação. Na região central do Brasil, identifica-se um expressivo potencial para a produção de trigo tropical (SILVA, 2022). Nesses territórios, os objetivos dos programas de melhoramento ampliam-se para contemplar o desenvolvimento de cultivares adaptadas a áreas irrigadas, além de apresentarem tolerância ao calor e à seca nos sistemas de produção em sequeiro.

2.4 Uso de Modelos Mistos

A avaliação de materiais genéticos busca inferir valores genotípicos e ordená-los para fins de seleção e/ou recomendação (RESENDE, 2007). Com a utilização de modelos mistos é possível estimar os componentes de variância e predizer valores genotípicos simultaneamente. Por meio da máxima verossimilhança residual (REML), estima-se a os componentes de variância, e a predição dos valores genotípicos é feita pela melhor predição linear não viesada (BLUP).

O BLUP é utilizado para a predição dos efeitos genéticos aditivos, de dominância e genotípicos. Com a utilização do REML é possível a decomposição da variância fenotípica total em vários componentes, genético, residual, ambiental e outro (RESENDE, 2007). Assim, torna-se viável a estimativa de parâmetros genéticos, possibilitando a compreensão do controle genético da característica de interesse. Isso visa estabelecer estratégias eficazes para o programa de melhoramento e a predição dos valores genotípicos dos candidatos à seleção.

De acordo com Resende (2016) as principais vantagens do uso de modelos mistos é estimar os parâmetros e predizer os valores genotípicos; viabiliza a comparação de indivíduos ao longo do tempo (anos) e em diferentes locais (espaço); aumenta a acurácia seletiva; otimiza o ganho genético e a eficiência dos programas de melhoramento; não exige dados balanceados; permite utilizar simultaneamente grande número de dados, gerando estimativas precisas; permite lidar com estruturas complexas de dados e permite a inclusão de informações de parentesco dentro do modelo.

Mediante as principais vantagens de se trabalhar com modelos mistos, existe uma série de trabalhos que implementaram essa abordagem estatística para estimar com mais precisão os parâmetros genéticos e residuais. Mezzomo et al. (2021) utilizando o método Jinks e Pooni (1976), utilizando as variâncias genéticas via REML e os BLUPs concluíram que é uma alternativa eficiente para seleção de populações superiores. Assim, como outros trabalhos Lima et al. (2022), Zhou et al. (2022) com trigo, especificamente.

3. MATERIAL E MÉTODOS

3.1 Material genético e condução experimental

Foram avaliadas em campo 20 populações de trigo nas safras de inverno e verão de 2021 a 2023. As populações foram obtidas a partir de um dialelo parcial (5x5), proveniente de 10 cultivares recomendadas para cultivo na região do Cerrado brasileiro. Os cultivares utilizados como genitores foram: Tbio Audaz, Tbio Astro, Tbio Aton, Tbio Duque e Tbio Toruk (Biotrigo Genética) no primeiro grupo e ORS Madre Pérola, ORS Citrino e ORS Destak (OR genética de Sementes), RBO 2B5 e RBO 3B6 (Tamona Agropecuária Ltda) no segundo grupo, como é apresentado na Tabela 1, com suas respectivas características.

Tabela 1 – Características agrônômicas de cultivares utilizados em cruzamento dialélico parcial em trigo

Cultivar	Ciclo	Altura	Doenças	Classe
ORS Citrino	P	Média	MR	Melhorador
ORS Destak	M	Média	R	Melhorador
ORS Madrepérola	P	Média	MR	Branqueador
RBO 2B5	P	Média	R	Branqueador
RBO 2B6	M	Baixa	R	Pão
Tbio Astro	SP	Baixa	MR	Melhorador
Tbio Aton	M	Média	R	Pão
Tbio Duque	P	Média	R	Branqueador
Tbio Toruk	M	Baixa	MR	Melhorador

M: médio; P: precoce; SP: superprecoce; MR: moderadamente resistente; R: resistente.

As populações nas gerações F₂, F₃ e F₄ foram avaliadas na área experimental Professor Diogo Alves de Melo (20° 45' 14" S, 42° 52' 55" W, 648 m) do Departamento de Agronomia da Universidade Federal de Viçosa, MG, Brasil, e somente a geração F₄, na Estação experimental Rio Paranaíba (19° 12' 26" S, 46° 09' 46" W, 1194 m), pertencente à Cooperativa Agropecuária do Alto Paranaíba (COOPADAP), São Gotardo, MG, Brasil em condições de sequeiro e irrigado.

Os experimentos foram organizados em delineamento em blocos ao acaso (DBC), com duas (São Gotardo) e três (Viçosa) repetições. As parcelas experimentais

foram constituídas por três linhas de 3 m, espaçadas a 0,20 m. As sementes foram semeadas por metro, seguindo o método de pedigree, proposto por MCVeetty e Evans (1980). Os experimentos foram conduzidos em condições de sequeiro e irrigado. Para atender às necessidades hídricas da cultura durante o desenvolvimento, os experimentos foram conduzidos com irrigação por aspersão.

3.2 Análises estatísticas

As análises foram conduzidas empregando a metodologia de modelos lineares mistos (REML/BLUP).

3.2.1 Modelo para as análises individuais

O modelo estatístico utilizado para detecção da variabilidade genética entre populações, foi o seguinte:

$$y = X\mu + Zg + \varepsilon$$

Onde y é o vetor de observações fenotípicas; μ é o vetor de efeitos fixos adicionado à média geral; g é o vetor de efeitos aleatórios dos genótipos, $g \sim N(0; \sigma_g^2)$; ε é o vetor de efeitos aleatórios dos resíduos, $\varepsilon \sim N(0; \sigma_\varepsilon^2)$. X , e Z são as matrizes de incidência para esses efeitos.

3.2.2 Modelo para a análise conjunta

O modelo estatístico para análise conjunta, utilizado para detectar os efeitos das populações e as interações populações x ambientes, foi o seguinte:

$$y = X_1\mu + X_2e + Zg + Wa + \varepsilon$$

Onde y é o vetor de observações fenotípicas; μ é o vetor de efeitos fixos de blocos adicionado à média geral; e é o vetor de efeitos fixos do ambiente; g é o vetor de efeitos aleatórios dos genótipos, $g \sim N(0; \sigma_g^2)$; a é o efeito aleatório da interação entre

genótipos e ambientes, $\mathbf{a} \sim N(0; \sigma_a^2)$; $\boldsymbol{\varepsilon}$ é o vetor de efeitos aleatórios dos resíduos, $\boldsymbol{\varepsilon} \sim N(0; \sigma_\varepsilon^2)$. \mathbf{X}_1 , \mathbf{X}_2 , \mathbf{Z} e \mathbf{W} são as matrizes de incidência para esses efeitos.

3.2.3 Teste de significância dos efeitos aleatórios

A significância dos efeitos aleatórios foi testada utilizando o teste da razão de verossimilhança (LRT) (WILKS, 1938) da seguinte forma:

$$LRT = -2 (\text{Log}L_F - \text{Log}L_R)$$

onde $\text{Log}L_F$ é o logaritmo da função de verossimilhança restrita do modelo completo, e $\text{Log}L_R$ é o logaritmo da função de verossimilhança restrita do modelo reduzido. A significância dos efeitos aleatórios foi testada pelo teste qui-quadrado a uma probabilidade de 5%.

3.2.4 Parâmetros genéticos

A partir dos modelos foram estimados os componentes de variância associados, coeficiente de variação experimental (CV) (%) e herdabilidade no sentido amplo (H^2).

$$CV (\%) = \frac{100 * \sqrt{\hat{\sigma}_\varepsilon^2}}{\mu}$$

onde σ_ε^2 representa a variância residual e μ representa a média fenotípica geral.

A herdabilidade a nível de média, em sentido amplo foi estimada da seguinte forma:

$$H^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_\varepsilon^2}$$

onde $\hat{\sigma}_g^2$ representa a variância genotípica das populações e $\hat{\sigma}_\varepsilon^2$ representa a variância residual.

3.2.5 Seleção de modelo

Para a análise conjunta, foi realizada a escolha de modelos com diferentes estruturas de (co)variância. Primeiramente, o efeito genético foi testado considerando estruturas de identidade de variância (IDV), estrutura diagonal (DIAG), estrutura de simetria composta (CS), simetria composta heterogênea (CSH) e não-estruturada (US). Após a escolha do modelo que mais se ajustou, considerando estrutura genética, foram modelados os efeitos residuais com a estrutura diagonal (DIAG).

A estrutura de IDV é homogênea, os elementos da diagonal apresentam o mesmo valor positivo. Nesse caso, há uma variância homogênea e covariância nula entre ambientes.

$$IDV = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 \\ 0 & 0 & 0 & \sigma_1^2 \end{bmatrix}$$

onde $\sigma_1^2 =$ variância.

A estrutura DIAG é heterogênea, os elementos da diagonal principal são positivos, no entanto, divergentes. Nessa estrutura há uma variância heterogênea nos ambientes e covariâncias nulas entre os pares de ambientes.

$$DIAG = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

A estrutura CS apresenta na sua diagonal a soma de duas variâncias distintas, porém, homogênea em todos os ambientes e a covariância homogênea para todos os pares de ambientes.

$$CS = \begin{bmatrix} \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 \end{bmatrix}$$

A estrutura de CSH considera parâmetros de variância heterogênea para cada ambiente e raiz quadrada desses parâmetros nos elementos fora da diagonal principal.

$$CSH = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2p & \sigma_1\sigma_3p & \sigma_1\sigma_4p \\ \sigma_2\sigma_1p & \sigma_2^2 & \sigma_2\sigma_3p & \sigma_2\sigma_4p \\ \sigma_3\sigma_1p & \sigma_3\sigma_2p & \sigma_3^2 & \sigma_3\sigma_4p \\ \sigma_4\sigma_1p & \sigma_4\sigma_2p & \sigma_4\sigma_3p & \sigma_4^2 \end{bmatrix}$$

Onde p é o parâmetro de correlação.

A estrutura US apresenta variância genética heterogênea para cada ambiente e covariância distinta entre os pares de ambientes.

$$US = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

Para seleção dos melhores modelos, para cada característica, foi realizada utilizando o critério de informação de Akaike (AIC, AKAIKE, 1974) e o critério de informação Bayesiano (BIC; SCHWARZ, 1978), fornecidos, respectivamente:

$$AIC = -2\log L + 2p$$

$$BIC = -2\log L + p\text{Log}[n - r(x)]$$

onde $\log L$ é o logaritmo da função de verossimilhança restrita, p é o número de parâmetros estimados, n é o número de observações, e $r(x)$ é a classificação da matriz de incidência de efeitos fixos (WOLFINGER, 1993).

3.2.6 Coeficiente de coincidência

O índice de coincidência (HAMBLIN E ZIMMERMANN, 1986) entre os ranks das gerações foi calculado da seguinte forma:

$$CI = \frac{A - C}{M - C} \times 100$$

Onde *A* é o número de genótipos coincidentes; *C* é o número de genótipos esperados, selecionados ao acaso, e *M* é o número de genótipos selecionados de acordo com a intensidade de seleção.

Todas as análises estatísticas foram conduzida no ambiente estatístico R, versão 4.3.1 (R CORE TEAM, 2023). Para a modelagem de efeitos mistos, utilizou-se os pacotes ASReml, versão 4.2.0.267 (THE VSNI TEAM, 2023) e ASRemlPlus, versão 4.4.21 (BRIEN, 2023).

4. RESULTADOS E DISCUSSÃO

O resumo das análises de deviance realizadas em diversas gerações, com base em dados fenotípicos de produtividade (PROD, ton ha⁻¹), altura de planta (ALT, cm) e espigamento (ESP), está apresentado na Tabela 2. Os resultados revelaram efeito significativo ($p \leq 0,05$) pelo teste LRT na geração F₂ para todas as variáveis analisadas, indicando a presença de variabilidade genética entre as populações. Na geração F₃, observou-se diferença significativa ($p \leq 0,05$) para todas as variáveis analisadas, com exceção de PROD. Na geração F₄, verificou-se um efeito significativo ($p \leq 0,05$) para todas as variáveis avaliadas em todos os ambientes, exceto para a variável PROD em São Gotardo sequeiro (SG_S).

A significância do efeito genotípico indica a presença de variabilidade genética entre as populações, um aspecto fundamental que pode ser explorado pelo melhorista por meio da seleção. A variabilidade genética é condição essencial para orientar programas de melhoramento. Ao compreender a variação genotípica dentro das populações, o melhorista pode direcionar o programa de acordo com os objetivos pré-estabelecidos (LIMA et al., 2022), identificando genótipos superiores com atributos agrônômicos favoráveis para alcançar maiores ganhos genéticos.

Tabela 2 - Análise da razão de verossimilhança (LRT) individual para todos os ambientes

Caráter	Geração F ₂		Geração F ₃		Geração F ₄	
	VI_S	VI_I	VI_S	VI_I	SG_S	SG_I
PROD	5,94**	1,81 ^{NS}	35,07**	17,30**	1,57 ^{NS}	9,48**
ALT	13,68**	7,75**	4,88*	10,68**	5,39*	4,47*
ESP	7,89**	52,49**	10,84**	57,95**	-	-

*, ** significativo a 5 e 1% de probabilidade, respectivamente pelo teste LRT. ^{NS}: Não significativo. PROD: produtividade; ALT: altura de planta e ESP: espigamento; VI_S: Viçosa sequeiro; VI_I: Viçosa irrigado; SG_S: São Gotardo sequeiro e SG_I: São Gotardo irrigado.

A partir do desdobramento da variável PROD, observou-se que os coeficientes de variação apresentaram valores significativamente elevados na maioria dos ambientes, variando de 28% a 8,75% (Tabela 3). Nardino et al. (2023), com o objetivo de estabelecer intervalos de confiança representativos dos coeficientes de variação (CV%) para a cultura do trigo, estudaram a distribuição de um conjunto de dados em consonância com a estatística descritiva, empregando meta-análise por meio de abordagens frequentista e bayesiana. Como resultado, concluíram que valores de CV% acima de 13% estão fora do intervalo de confiança para a produtividade.

Tabela 3 - Estimativa dos componentes de variâncias nos respectivos ambientes

Variável	Ambiente A VI_S F2					
	σ_e	σ_f	σ_g	H^2	CVe (%)	Média
PROD	0,65	1,15	0,50	0,60	26,18	3,08
ALT	19,43	50,25	30,82	0,76	4,99	88,41
ESP	0,92	1,80	0,88	0,66	1,42	67,56
Variável	Ambiente B VI_I F3					
	σ_e	σ_f	σ_g	H^2	CVe (%)	Média
PROD	1,72	2,13	0,41	0,42	19,06	6,88
ALT	34,48	57,34	22,86	0,67	5,76	101,91
ESP	1,50	10,66	9,16	0,95	2,03	60,24
Variável	Ambiente C VI_S F4					
	σ_e	σ_f	σ_g	H^2	CVe (%)	Média
PROD	0,04	0,13	0,09	0,88	23,14	0,82
ALT	19,93	28,20	8,27	0,55	6,47	69,04
ESP	15,14	31,03	15,89	0,76	7,21	54,00
Variável	Ambiente D VI_I F4					
	σ_e	σ_f	σ_g	H^2	CVe (%)	Média
PROD	0,18	0,39	0,21	0,78	28,00	1,52

ALT	12,66	22,29	9,63	0,69	4,02	88,42
ESP	1,31	8,58	7,27	0,94	1,98	57,68
Ambiente E SG_S F4						
Variável	σ_e	σ_f	σ_g	H^2	CVe (%)	Média
PROD	0,17	0,33	0,16	0,65	10,29	5,38
ALT	6,21	19,37	13,16	0,81	5,72	80,76
ESP	-	-	-	-	-	-
Ambiente F SG_I F4						
Variável	σ_e	σ_f	σ_g	H^2	CVe (%)	Média
PROD	0,46	0,79	0,33	0,69	8,75	7,75
ALT	83,26	116,60	33,34	0,55	8,77	104,03
ESP	-	-	-	-	-	69,87

Efeito residual (σ_e); componente de variância fenotípica (σ_f); componente de variância genotípica (σ_g); herdabilidade (H^2); coeficiente de variação (CVe%); PROD: produtividade de grãos; ALT: altura de planta e ESP: espigamento; VI_S: Viçosa sequeiro; VI_I: Viçosa irrigado; SG_S: São Gotardo sequeiro; SG_I: São Gotardo irrigado e Média: média fenotípica.

Os coeficientes de variação mais elevados foram identificados especificamente para PROD, o que era esperado, devido ao seu caráter poligênico. Dado que o coeficiente de variação (CVe%) é a relação do desvio padrão em relação à média, a natureza poligênica da variável PROD, aliada ao trabalho com populações segregantes, contribui para a obtenção de valores elevados. Para as variáveis ALT e ESP, todos os coeficientes de variação (CVe%) estão adequados, conforme destacado no estudo de Nardino et al. (2023), apresentando variações de 8,77% a 4,02% para ALT e de 7,21% a 1,42% para ESP.

A herdabilidade apresentou valores significativamente divergentes em relação ao caráter e aos ambientes. No entanto, conforme apontado por Amaral et al. (1996), valores diferenciados de herdabilidade não são exclusivamente uma característica intrínseca do caráter, mas dependem da variância genética das populações e das variâncias dos ambientes em que essas populações estão inseridas. Nesse contexto, para as variáveis em que a participação do ambiente foi mínima, justificam-se valores mais elevados de herdabilidade. Dessa forma verifica-se a alta herdabilidade para os caracteres avaliados (0,55 a 0,95%), exceto PROD no ambiente B VI_I F₃ (0,42%). Indicando que há possibilidade de ganho com a seleção.

Os ensaios com populações segregantes, testados em diferentes gerações, são afetados de maneira distinta em cada ambiente, devido à influência de diversos fatores, especialmente quando se trata de uma característica poligênica. Mesmo entre populações da mesma geração, observa-se uma mudança significativa na

característica. Na Tabela 3, verifica-se que a média de PROD variou de 0,82 a 7,75 ton ha⁻¹, resultante de vários fatores, como a constituição genotípica, ambiental, fisiológica, entre outros. Portanto, de acordo com Chaves et al. (2023), é essencial considerar a covariância e a heterogeneidade no modelo estatístico utilizado para a análise de dados.

O resumo das estimativas dos critérios de informação de Akaike (AIC) e Schwarz ou Bayesiano (BIC), empregados na seleção de modelos para a análise conjunta considerando efeitos genotípicos e residuais com base em dados fenotípicos de produtividade, é apresentado na Tabela 4. Este critério se baseia em eficiência, isto é, o melhor modelo é aquele que minimiza a variância do erro (CASELLA e BERGHER, 2002). Como o objetivo é minimizar a perda de informações, o modelo com menor valor (em pelo menos 2 unidades) AIC e BIC tem maior suporte nos dados (RESENDE et al., 2020; NEATH e CAVANAUGH, 2012).

Segundo os critérios AIC e BIC, a estrutura de CSH (M4) foi a mais adequada para explicar os efeitos genéticos, apresentando valores de AIC e BIC de 259,42 e 278,31, respectivamente. Quando se considera o modelo CSH para covariância genotípica, modelando também o efeito residual com DIAG (M6), esses valores dos critérios de seleção de modelo diminuem para 186,68 e 216,90, respectivamente. Isso representa uma diferença de 80,4 (AIC) e 61,5 (BIC) em relação ao modelo que considera as estruturas homogêneas (IDV). Ambos os critérios de seleção, em consonância com os parâmetros genéticos e residuais, indicam a superioridade do modelo que considera as variâncias heterogêneas em detrimento ao modelo que as considera homogêneas.

Esses resultados demonstram a importância de considerar a heterogeneidade das variâncias, testando diversas estruturas de covariância com diferentes níveis de complexidade. Em todos os modelos em que apenas a covariância genotípica foi considerada, houve oscilação nos valores de AIC e BIC, sendo sempre iguais ou superiores ao IDV. No entanto, uma vez identificada a melhor estrutura para os efeitos genotípicos e modelando o efeito residual, alcançou-se uma maior precisão nos modelos (M6, Tabela 4).

Tabela 4 - Estrutura de covariância genotípica (Σ_g) e efeito residual (Σ_e), modelo pelos critérios de informação de Akaike (AIC) e de Schwarz ou bayesiano (BIC), logaritmo

da função de verossimilhança ($\log(L)$), número de parâmetros associados a Σ_g ($\Sigma_g \text{ par}$) e Σ_e ($\Sigma_e \text{ par}$) para a variável produtividade de grãos (PROD)

Model	Σ_g	Σ_e	AIC	BIC	$\log(L)$	$\Sigma_g \text{ par}$	$\Sigma_e \text{ par}$
M1	IDV	IDV	267,13	278,46	-132,00	3	1
M2	DIAG	IDV	260,76	275,87	-126,38	4	1
M3	CS	IDV	267,13	278,46	-130,56	3	1
M4	CSH	IDV	259,42	278,31	-124,71	5	1
M5	US	IDV	263,17	289,62	-124,58	7	1
M6	CSH	DIAG	186,68	216,90	-85,34	5	3

ID: matriz de identidade de variâncias; DIAG: diagonal; CS: simetria composta; CSH: simetria composta heterogênea; US: matriz não estruturada.

Dada a escolha do modelo mais apropriado para as estimativas de componentes genéticos e residuais, foram estimados os BLUPs das populações, conforme mostrado na Tabela 4. O coeficiente de coincidência entre as gerações F_2 e F_3 , na seleção das 10 populações de melhor desempenho (pressão de seleção de 50%), foi de 0,80; entre as gerações F_2 e F_4 ; F_3 e F_4 , foi de 0,60. Esses resultados indicam um certo grau de similaridade entre as populações selecionadas precocemente nas gerações F_2 em relação a F_4 .

Tabela 5 – Média dos BLUPs das 20 populações em três gerações utilizadas neste estudo nas safras 2021, 2022 e 2023 para produtividade de grãos (PROD, ton ha⁻¹) utilizando o modelo (M6) de simetria composta heterogênea (CSH) para efeito genotípico e diagonal (DIAG) para efeito residual

CULTIVARES	F_2	F_3	F_4	MÉDIA
TBIO ASTRO	4.38	7.49	3.91	5.26
TBIO AUDAZ	3.48	7.07	3.67	4.74
ORS DESTAK	3.21	7.10	3.86	4.72
TBIO ATON	3.08	7.04	3.83	4.65
ORS CITRINO	3.26	6.97	3.62	4.62
TBIO DUQUE	3.00	6.85	3.55	4.47
TBIO TORUK	3.00	6.85	3.55	4.46
ORS MADRE PÉROLA	2.85	6.78	3.51	4.38
RBO 2B5	2.73	6.72	3.48	4.31
RBO 3B6	2.21	6.48	3.34	4.01
POPULAÇÕES	F_2	F_3	F_4	MÉDIA
TBIO ATON / ORS DESTAK	3.61	7.03	4.14	4.93
TBIO ATON / ORS CITRINO	3.14	7.60	3.89	4.88
TBIO DUQUE / ORS CITRINO	3.51	7.24	3.86	4.87
TBIO AUDAZ / ORS CITRINO	3.85	7.09	3.60	4.85
TBIO ATON / RBO 2B5	3.40	7.35	3.72	4.82

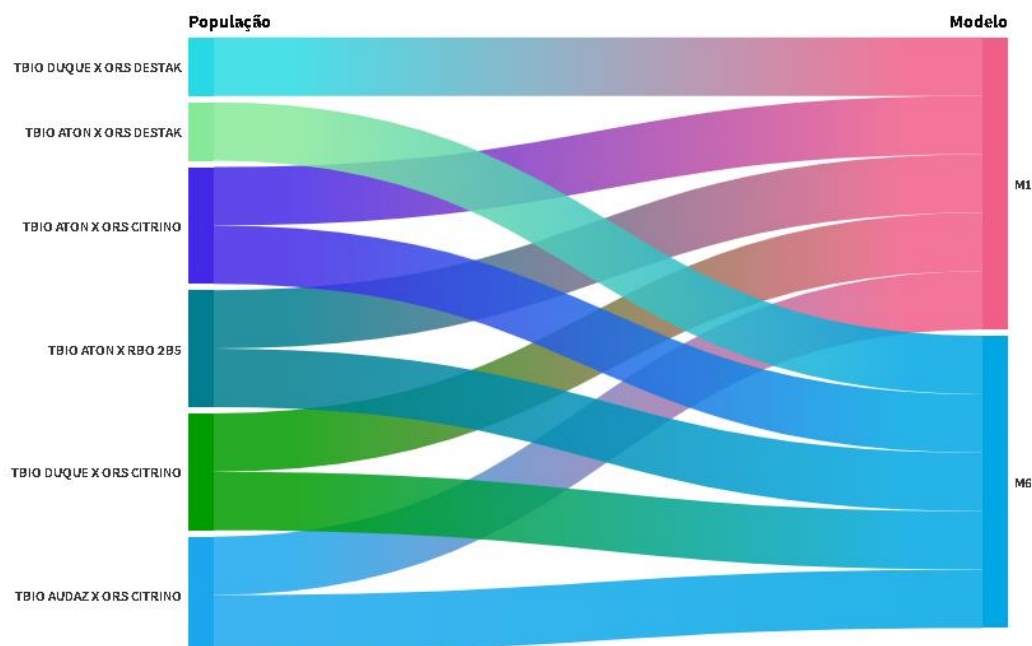
TBIO DUQUE / ORS DESTAK	3.41	7.17	3.86	4.81
TBIO TORUK / ORS CITRINO	3.45	7.13	3.28	4.62
TBIO TORUK / ORS DESTAK	3.41	7.05	3.38	4.62
TBIO ATON / ORS MADRE PÉROLA	3.08	7.05	3.68	4.60
TBIO DUQUE / RBO 2B5	3.28	6.90	3.50	4.56
TBIO ASTRO / ORS DESTAK	2.94	6.94	3.74	4.54
TBIO ASTRO / ORS MADRE PÉROLA	3.04	6.90	3.50	4.48
TBIO DUQUE / ORS MADRE PÉROLA	2.75	6.64	3.60	4.33
TBIO ATON / RBO 3B6	2.62	6.89	3.42	4.31
TBIO TORUK / RBO 2B5	2.81	6.65	3.45	4.30
TBIO ASTRO / ORS CITRINO	2.89	6.60	3.41	4.30
TBIO ASTRO / RBO 2B5	2.80	6.24	3.30	4.11
TBIO TORUK / ORS MADRE PÉROLA	2.61	6.46	3.04	4.04
TBIO ASTRO / RBO 3B6	2.43	6.27	3.31	4.00
TBIO TORUK / RBO 3B6	2.26	6.07	3.08	3.80
IC F2:F3	0.8			
IC F2:F4	0.6			
IC F3:F4	0.6			

IC: Índice de coincidência

O índice de coincidência entre populações F_2 e F_3 (0,80) foi maior que para F_2 e F_4 ; F_3 e F_4 (0,60). No entanto, é interessante notar que todas as populações ranqueadas em F_2 são as mesmas ranqueadas com base na média dos BLUPs das gerações. Isso demonstra que as populações selecionadas em estágios precoces apresentam desempenho consistente ao longo das gerações. Considerando isso, podemos utilizar a seleção em fases iniciais de segregação, reduzindo tempo e economizando recursos.

A variabilidade genética é uma condição essencial para direcionar um programa de melhoramento de trigo. Compreendendo a variação genotípica dentro das populações, o melhorista pode direcionar o programa conforme os objetivos pretendidos (FAYE et al., 2021; LIMA et al., 2021). No entanto, essa variação deve conter superioridade das populações em relação aos pais. Na Tabela 4, é possível observar que, em média, nas dez melhores populações, os BLUPs foram iguais ou superiores a pelo menos um dos progenitores.

Figura 1 – Populações selecionadas para características de produtividade usando dois modelos: M1, matriz de identidade de variâncias e M6, combinação da simetria composta heterogênea para efeito genético e matriz diagonal para efeito residual



Baseado na média dos BLUPs das populações, estimados pelo modelo M6, a melhor população foi a Tbio Aton X ORS Destak, no entanto, como é apresentado na Figura 1, essa população sequer foi selecionada, quando comparamos ao modelo M1. Uma vez que afeta o rank das populações e que pode afetar diretamente a decisão do melhorista, o uso de modelos estatísticos mais precisos, aumenta a probabilidade de levar adiante materiais superiores.

A melhor cultivar, baseados nos valores de BLUPs, é a Tbio Astro. No entanto, não gerou as melhores populações. De acordo com Borém et al. (2021), quando características indesejáveis são ligadas a características desejáveis, a probabilidade de obtenção de indivíduos superiores depende de *crossing overs* específicos para uma maior probabilidade de obtenção de populações segregantes com bons parâmetros genéticos. Outra questão que pode ser levada em consideração em relação as melhores cultivares pode ser atribuída aos desvios de dominâncias.

7. CONCLUSÕES

O melhor modelo com diferentes estruturas de modelagem das (co)variâncias residuais e genéticas para uma análise de modelos mistos foi a combinação de simetria composta heterogênea para efeito genéticos, e de simetria diagonal para efeito residual (M6).

Com base nos resultados, a seleção de populações de melhor performance pode ser realizada nas gerações iniciais do programa de melhoramento, tais como F₂.

As melhores populações com base no modelo mais apropriado para as estimativas dos BLUPs foram as TBIO ATON / ORS DESTAK, TBIO ATON / ORS CITRINO, TBIO DUQUE / ORS CITRINO, TBIO AUDAZ / ORS CITRINO e TBIO ATON / RBO 2B5.

8. REFERÊNCIAS

- ALI, Z. et al. Wetting mechanism and morphological adaptation; leaf rolling enhancing atmospheric water acquisition in wheat crop — a review. **Environ Science Pollut Res** v. 29, p. 30967–30985, 2022.
- AKAIKE, H. **Information theory and an extension of the maximum likelihood principle** (2nd ed.). BN Petrov, 1973.
- AMARAL, A.L. et al. Estimativa da herdabilidade para os caracteres adaptativos ciclo e estaturas de planta em aveia. **Ciência Rural**, v. 26 (1), p. 33-37, 1996.
- BRIEN, C. (2023). `_asremlPlus`: Augments 'ASReml-R' in Fitting Mixed Models and Packages Generally in Exploring Prediction Differences_. R package version 4.4.21, <<https://CRAN.R-project.org/package=asremlPlus>>.
- BORÉM, A.; SCHEEREN, P.L. **Trigo do plantio à colheita**. UFV: Viçosa, 2015. p. 11-34.
- BORÉM, A; MIRANDA, G.V.; FRITSCHÉ-NETO, R. **Melhoramento de plantas**. Oficina de Textos, 2021.
- CASELLA G.; BERGER R. (2002). **Statistical Inference**, 2nd ed. Duxbury, Thomson Learning, CA.
- CHAVES, S.F.S. et al. Analysis of repeated measures data through mixed models: An application in *Theobroma grandiflorum* breeding. **Crop Science**, v. 63, p. 2131–2144, 2023.
- CONAB. Observatório agrícola. **Acompanhamento da safra brasileira de grãos**. 5^º levantamento. Fevereiro de 2024. Disponível em: https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos/item/download/51814_c19d452a0b14ed1e74b0d45ae0b815e5. Acessado em: 15 fevereiro de 2024.
- DIAS, K.O.G., et al. Novel strategies for genomic prediction of untested singlecross maize hybrids using unbalanced historical data. **Theoretical and Applied Genetics**, v. 133, p. 443–455, 2020.
- EGEA-GILABERT, C; PAGNOTTA, M.A.; TRIPODI, P. Genotype × Environment Interactions in Crop Breeding. **Agronomy**, v. 11(8). p. 1644, 2021.
- FALCONER, D.S.; MACKAY, T.F. **Introduction to Quantitative Genetics**, 4th ed. Longman, Harlow, 1996.
- FAYE, J.M. et al. A genomics resource for genetics, physiology, and breeding of West African sorghum. **Plant Genome**, v. 14: e20075, 2021.

- FELDMAN, M.; LUPTON, F.G.H.; MILLER, T.E. Wheats. In: SMARTT, J.; SIMMONDS, N. W. **Evolution of crop plants**. Harlow: Longman, 1995. p. 184-192.
- GEZAN, S. A.; DE CARVALHO, M. P.; SHERRILL, J. Statistical methods to explore genotype-by-environment interaction for loblolly pine clonal trials. **Tree Genetics & Genomes**, v.13(1), p. 1–11, 2017.
- HALL, ANTONIO J.; RICHARDS, RICHARD A. Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. **Field Crops Research**, v. 143, p. 18-33, 2013.
- HÄMÄLÄ, T. et al. Environmental response in gene expression and DNA methylation reveals factors influencing the adaptive potential of *Arabidopsis lyrata*. **Elife**, v.11. e831152022.
- HAMBLIN, J.E.; ZIMMERMANN, M.J.O. Breeding common bean for yield in mixtures. **Plant Breeding Reviews**, v.4, p.245-272. 1986.
- HUANG, S. et al. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/ Aegilops* complex and the evolutionary history of polyploid wheat. **Proceedings of the National Academy of Sciences of the United States of America**, v. 99, p. 8.133-8.138, 2002.
- HUDSON, A.I. et al. Analysis of genotype-by-environment interactions in a maize mapping population. **G3 (Bethesda)**, v. 12. 20909, 2022.
- JINKS, J.L. POONI, H.S. Predicting the properties of recombinant inbred lines derived single seed descent. **Heredity**, v. 36, p. 243-266, 1976.
- JÚNIOR, R.S. et al. Achieving wheat self-sufficiency in Brazil. **Environmental Research Letters**, v. 19, p. 031003, 2024.
- LEE, S.Y. et al. Multi-Environment Trials and Stability Analysis for Yield-Related Traits of Commercial Rice Cultivars. **Agriculture**, v.13, p. 256, 2023.
- LI, RONGHUA, et al. Exploring the relationships between yield and yield-related traits for rice varieties released in China from 1978 to 2017. **Frontiers in plant science**, v. 10, p. 543, 2019.
- LIMA, G.W. et al. Genetic diversity in tropical wheat germplasm and selection via multitrait index. **Agronomy Journal**, v. 114, n. 2, p. 887-899, 2022.
- LIMA, G.W. et al. Canonical correlations between high and low heritability wheat traits via mixed models. **Ciência Rural**, v. 53, p. e20210798, 2022.
- MALOSETTI, M.; RIBAUT, J.M.; EEUWIJK, F.A.V. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. **Frontiers in Physiology**. v.4, p. 1-17, 2013.
- MCVETTY P.B.E.; EVANS L.E. Breeding methodology in wheat. II. Productivity

harvest index, and height measured on F2 spaced plants for yield selection in spring wheat. **Crop Science** v. 20: p. 587-589, 1980.

MEZZOMO, H.C. et al. Mixed model-based Jinks and Pooni method to predict segregating populations in wheat breeding. **Crop Breeding and Applied Biotechnology**, v. 21, p. e38352143, 2021.

NARDINO, M. et al. Meta-analysis of the experimental coefficient of variation in wheat using the Bayesian and Frequentist approaches. **Scientia Agricola**, v. 80, p. e20210190, 2022.

NEATH, A.A.; CAVANAUGH, J.E. The Bayesian information criterion: background, derivation, and applications. **Computational Statistics**, v. 4, p.199-203, 2012.

PENG, S. et al. Yield potential trends of tropical rice since the release of IR8 and the challenge of increasing rice yield potential. **Crop Science**, v. 39, n. 6, p. 1552-1559, 1999.

PIANA, C.F.B.; CARVALHO, F.I.F. Trigo. In BARBIERI, R.L.; STUMPF, E.R.T. **Origem e evolução das plantas cultivadas**. Brasília: Embrapa, 2008, p. 821-852.

R CORE TEAM (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>

REIS, R.R. et al. Differential gene expression in response to water deficit in leaf and root tissues of soybean genotypes with contrasting tolerance profiles. **Genetics and Molecular Biology**, v. 43, e20180290, 2020.

RESENDE, M.D.V. **Matemática e estatística na análise de experimentos e no melhoramento genético**, 1st ed. Embrapa (ed). Embrapa Florestas, Colombo, 561p, 2007.

RESENDE, M.D.V. Software Selegen-REML/BLUP: A useful tool for plant breeding. **Crop Breeding and Applied Biotechnology** 16: 330–339, 2016.

RESENDE, M.D.V.; ALVES, R.S. Linear, generalized, hierarchical, bayesian and random regression mixed models in genetic/genomics in plant breeding. **Functional Plant Breeding Journal**, v. 2, n. 2, p. 1–31, 2020.

ROGERS, A.R. et al. The importance of dominance and genotype-by-environment interactions on grain yield variation in a large-scale public cooperative maize experiment. **G3 (Bethesda)**, v. 11 (2), 2021.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, 6, 461-464. 1978.

SINGH, M. et al. Male fertility genes in bread wheat (*Triticum aestivum* L.) and their utilization for hybrid seed production. **International Journal of Molecular Sciences**, v. 22, n. 15, p. 8157, 2021.

SILVA, C.M. Seleção de genitores e potencial de populações segregantes de trigo. 2022. 67 f. **Dissertação (Mestrado em Genética e Melhoramento)** - Universidade Federal de Viçosa, Viçosa. 2022.

SMITH, A. B.; CULLIS, B. R.; THOMPSON, R. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. **Journal of Agricultural Science**, v.143, p. 449– 462, 2005.

SOUSA, C.N.A.; ROSA, O.S. Multiplicar o grão. **Ciência Hoje**, v. 3, p. 46-52, 1985.

THE VSNI TEAM (2023). `_asreml: Fits Linear Mixed Models using REML_`. R package version 4.2.0.267, <www.vsni.co.uk>.

UCHENDU, K. et al. Genotype-by-environment interaction and stability of root mealiness and other organoleptic properties of boiled cassava roots. **Scientific Reports**, v.12, p. 20909, 2022.

USDA, 2024. Grain: World Markets and Trade. Disponível em: <<https://www.ers.usda.gov/data-products/wheat-data/>>. Acessado em 15 fevereiro de 2024.

VAN EEUWIJK, F.A. et al. What should students in plant breeding know about the statistical aspects of genotype× environment interactions?. **Crop Science**, v. 56, n. 5, p. 2119-2140, 2016.

WILKS, S.S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. **Annals of Mathematical Statistics**, v.1, p. 60–62, 1938.

WOLFINGER, R. **Covariance structure selection in general mixed models. Communications in Statistics - Simulation and Computation**, v. 22, p.1079–1106, 1993.

ZHOU, X. et al. Using linear mixed-effects modeling to evaluate the impact of edaphic factors on spatial variation in winter wheat grain yield in Japanese consolidated paddy fields. **European Journal of Agronomy**, v. 133, p. 126447, 2022.