

**BRUNA STEPHANI DE PAULA**

**SELEÇÃO SIMULTÂNEA DE GENÓTIPOS DE SOJA DE DIFERENTES GRUPOS  
DE MATURIDADE BASEADA EM CARACTERÍSTICAS AGRONÔMICAS E DE  
CURVA DE SENESCÊNCIA**

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: Felipe Lopes da Silva

Coorientadores: Luiz Alexandre Peternelli  
Lucas de Paula Corrêdo  
Domingos Sárvio Magalhães

Valente

**VIÇOSA - MINAS GERAIS  
2024**

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

T

P324s  
2024  
Paula, Bruna Stephani de, 1995-  
Seleção simultânea de genótipos de soja de diferentes  
grupos de maturidade baseada em características agronômicas e  
de curva de senescência / Bruna Stephani de Paula. – Viçosa,  
MG, 2024.

1 dissertação eletrônica (54 f.): il. (algumas color.).

Orientador: Felipe Lopes da Silva.

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

Referências bibliográficas: f. 49-54.

DOI: <https://doi.org/10.47328/ufvbbt.2024.226>

Modo de acesso: World Wide Web.

1. *Glicine max*. 2. Soja -Melhoramento genético.  
3. Fenótipo. 4. Drone. 5. Soja - Seleção. 6. Senescência. I. Silva,  
Felipe Lopes da, 1981-. II. Universidade Federal de Viçosa.  
Departamento de Agronomia. Programa de Pós-Graduação em  
Genética e Melhoramento. III. Título.

CDD 22. ed. 633.342

Bibliotecário(a) responsável: Alice Regina Pinto Pires CRB-6/2523


**BRUNA STEPHANI DE PAULA**

**SELEÇÃO SIMULTÂNEA DE GENÓTIPOS DE SOJA DE DIFERENTES GRUPOS DE MATURIDADE BASEADA EM CARACTERÍSTICAS AGRONÔMICAS E DE CURVA DE SENESCÊNCIA**

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: 16 de fevereiro de 2024

Assentimento:

Documento assinado digitalmente  
 **BRUNA STEPHANI DE PAULA**  
Data: 05/07/2024 11:33:18-0300  
Verifique em <https://validar.it.gov.br>

---

**Bruna Stephani de Paula**  
Autor

Documento assinado digitalmente  
 **FELIPE LOPES DA SILVA**  
Data: 15/07/2024 15:05:25-0300  
Verifique em <https://validar.it.gov.br>

---

**Felipe Lopes da Silva**  
Orientador

*Às minhas avós Marli Antonia Machado (in  
memorian) e Oliete Maria de Araujo Paula.*

## **AGRADECIMENTOS**

A Deus e aos espíritos superiores de luz por terem me guiado, dado forças nos momentos de dificuldade e tranquilidade nos momentos mais conturbados.

A minha família, pelo apoio, confiança, paciência, amor e principalmente por não terem medido esforços ao longo dos anos para que este sonho se tornasse realidade.

À Universidade Federal de Viçosa, por meio do Programa de Pós-Graduação em Genética e Melhoramento pela oportunidade de realizar a pós-graduação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

Ao Prof. Felipe Lopes da Silva, pela orientação do trabalho e pelo conhecimento transmitido.

Aos professores Luiz Alexandre Peternelli, Lucas de Paula, Domingos Sárvio Magalhães Valente pela coorientação e ensinamentos.

Aos técnicos do Laboratório de Genética da Soja, pelo auxílio durante a condução dos experimentos.

A todos os amigos do Laboratório de Genética da Soja por toda a ajuda, dedicação e suporte ao longo do desenvolvimento do trabalho.

Aos alunos de pós-doutorado do Laboratório de Análises e Pesquisas em Estatística Aplicada por toda a ajuda e suporte.

A todos meus amigos pela amizade e apoio ao longo destes anos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

Minha gratidão a todos que contribuíram de forma direta ou indireta para o bom andamento deste trabalho.

## RESUMO

de Paula, Bruna Stephani, M.Sc., Universidade Federal de Viçosa, janeiro de 2024. **Seleção simultânea de genótipos de soja de diferentes grupos de maturidade baseada em características agronômicas e de curva de senescência.** Orientador: Felipe Lopes da Silva. Coorientadores: Luiz Alexandre Peternelli, Lucas de Paula Corrêdo, Domingos Sárvio Magalhães Valente.

A soja desempenha um papel crucial na agricultura global, sendo o Brasil o maior produtor mundial. O melhoramento genético tem suma importância para alavancar a produtividade desta cultura, para atender a demanda do mercado mundial que vem aumentando ao longo dos anos. Este estudo buscou identificar os ideótipos de soja de diferentes grupos de maturação ao incluir o estudo de características agronômicas e de curvas de senescência em progênies de soja, e selecionar os genótipos mais produtivos em cada grupo de maturação utilizando o índice MGIDI (Multitrait Genotype Ideotype Distance Index), a partir do ideótipo desenhado para cada grupo. Durante a safra 2022/23, dois ensaios de teste de progênies F<sub>3:5</sub> foram realizados na Universidade Federal de Viçosa, adotando o delineamento de blocos aumentados (DBA). As parcelas experimentais foram constituídas por três linhas com 1,5 m de comprimento, espaçadas em 0,50 m, com densidade populacional de 16 plantas por metro. Foram avaliados: produção, data de florescimento, data de maturação, peso de 100 sementes, altura de planta na maturação e acamamento. A curva de maturação foi traçada utilizando valores de reflectância do índice GLI ao longo dos dias após o plantio (DAP), ajustados por regressão LOESS. As variáveis extraídas da curva foram: área total abaixo da curva (AUC 1), o ponto de máximo GLI (GLIMáx), o número de dias até o ponto de máximo GLI (DAMáx), o número de dias do ponto de máximo GLI até a maturação (DDMáx) e a área em baixo da curva obtida a partir do ponto de máximo GLI até a maturação completa (AUC 2) e verde acumulado, calculado pela diferença o GLIMáx e o valor inicial do índice GLI. A correlação de Pearson entre as variáveis da curva de maturação. Foi utilizando o índice MGIDI, com diferentes combinações de variáveis agronômicas e da curva de maturação, além da seleção direta através do BLUP de produtividade, comparando-se sete estratégias de seleções. O índice MGIDI mostrou eficácia nas seleções 2 a 6, enquanto a seleção 2 foi menos eficiente na população 1. Na população 2, a seleção 3 foi menos eficaz. A coincidência entre os 10 melhores genótipos das seleções variou

significativamente, destacando a importância de considerar a data de maturação ao conduzir a seleção. Estudar o comportamento dos genótipos dentro dos grupos de maturação revelou-se promissor, possibilitando uma seleção mais precisa e eficiente. A análise de imagem demonstrou eficiência na extração de variáveis da curva de maturação, através da utilização do índice de vegetação GLI (Green Leaf Index), embora seja necessário realizar mais estudos para correlacionar as fases fenológicas da soja com essas variáveis de maneira mais precisa. A análise das diferenças médias entre grupos reforçou a importância de conduzir a seleção considerando a data de maturação, evitando beneficiar genótipos de um grupo específico. Em conclusão, a associação das características agrônômicas e características da curva de maturação foram eficientes para definir o ideótipo de cada grupo em cada população e selecionar genótipos superiores em cada grupo a partir do ideótipo definido, nas diferentes estratégias de seleção.

Palavras-chave: *Glycine max*, Melhoramento da soja, Ideótipo, Fenotipagem via drone, Seleção multivariada, MGIDI

## ABSTRACT

de Paula, Bruna Stephani, M.Sc., Universidade Federal de Viçosa, January, 2024. **Simultaneous selection of soybean genotypes from different maturity groups based on agronomic traits and senescence curves.** Adviser: Felipe Lopes da Silva. Co-advisers: Luiz Alexandre Peternelli, Lucas de Paula Corrêdo, Domingos Sárvio Magalhães Valente

Soybean plays a crucial role in global agriculture, with Brazil being the world's largest producer. Genetic improvement is of paramount importance to boost the productivity of this crop, to meet the increasing demand of the global market over the years. This study aimed to identify soybean ideotypes of different maturity groups by including the study of agronomic characteristics and senescence curves in soybean progenies, and selecting the most productive genotypes in each maturity group using the MGIDI index (Multitrait Genotype Ideotype Distance Index), based on the ideotype designed for each group. During the 2022/23 growing season, two  $F_{3:5}$  progeny trials were conducted at the Federal University of Viçosa, adopting the augmented block design (ABD). The experimental plots consisted of three rows 1.5 m long, spaced at 0.50 m, with a population density of 16 plants per meter. The following were evaluated: yield, flowering date, maturity date, 100-seed weight, plant height at maturity, and lodging. The maturity curve was plotted using reflectance values of the GLI index over days after planting (DAP), adjusted by LOESS regression. Variables extracted from the curve included: total area under the curve (AUC 1), the point of maximum GLI (GLIMax), the number of days to the point of maximum GLI (DAMax), the number of days from the point of maximum GLI to maturity (DDMax), and the area under the curve from the point of maximum GLI to complete maturity (AUC 2), and accumulated green, calculated by the difference between GLIMax and the initial value of the GLI index. Pearson correlation between maturity curve variables was performed. The MGIDI index, with different combinations of agronomic and maturity curve variables, was used along with direct selection through BLUP of productivity, comparing seven selection strategies. The MGIDI index showed efficacy in selections 2 through 6, while selection 2 was less efficient in population 1. In population 2, selection 3 was less effective. The coincidence among the top 10 genotypes of selections varied significantly, highlighting the importance of considering maturity date when conducting selection. Studying the behavior of genotypes within maturity groups proved promising, enabling more precise

and efficient selection. Image analysis demonstrated efficiency in extracting variables from the maturity curve using the GLI (Green Leaf Index), although further studies are needed to correlate soybean phenological stages with these variables more accurately. Analysis of mean differences between groups reinforced the importance of conducting selection considering maturity date, avoiding favoring genotypes from a specific group. In conclusion, the association of agronomic characteristics and maturity curve characteristics were effective in defining the ideotype of each group in each population and selecting superior genotypes in each group based on the defined ideotype, in different selection strategies.

Keywords: Glycine max, Soybean breeding, Ideotype, Drone-based phenotyping, Multivariate selection, MGIDI.

## SUMÁRIO

1. INTRODUÇÃO.....	10
2. MATERIAL E MÉTODOS .....	13
<b>2.1. Material vegetal, descrição do local e coleta de dados.....</b>	<b>13</b>
<b>2.2. Planejamento, execução dos voos e avaliação da precisão do georreferenciamento .....</b>	<b>15</b>
<b>2.3 Processamento e análise de imagem .....</b>	<b>17</b>
2.3.1 Obtenção de dados baseados na curva de maturação construída por meio do índice GLI.....	19
<b>2.4 Análises estatísticas .....</b>	<b>21</b>
2.4.1 Correlação entre as variáveis.....	22
2.4.2 Seleção dos genótipos .....	23
3. RESULTADOS E DISCUSSÃO .....	24
<b>3.1 Variáveis da curva de maturação.....</b>	<b>24</b>
<b>3.2 Variáveis agronômicas.....</b>	<b>28</b>
<b>3.3 Construção do ideótipo e estudo do comportamento dos genótipos por data de maturação.....</b>	<b>30</b>
3.3.1 População 1 .....	33
3.3.1.1 Grupo 1.....	33
3.3.1.2 Grupo 2.....	34
3.3.1.3 Grupo 3.....	35
3.3.1.4 Grupo 4.....	35
3.3.2 População 2.....	36
3.3.2.1 Grupo 1.....	36
3.3.2.2 Grupo 2.....	37
3.3.2.3 Grupo 3.....	37
3.3.2.4 Grupo 4.....	38
<b>3.4 Seleção de genótipos.....</b>	<b>38</b>
3.4.1 População 1 .....	39
3.4.2 População 2 .....	41
3.4.3 Ganho com seleção .....	43
3.4.3.1 População 1.....	45
3.4.3.2 População 2.....	46
4. CONCLUSÃO .....	48
5. REFERÊNCIAS .....	49

## 1. INTRODUÇÃO

A soja (*Glycine Max L.*) destaca-se como uma das culturas mais importantes do mundo, uma produção mundial de quase 400 milhões de toneladas na safra 2023 (USDA/PSD, 2024). Essa importância deve-se principalmente pelo fato dessa cultura ser uma fonte de proteína e óleo vegetal (SILVA, et al. 2022), de uso amplo para alimentação animal e crescente na alimentação humana, além de ser utilizada como alternativa para fabricação de biocombustíveis.

Desde 2020, o Brasil é o maior produtor de soja do mundo, atingindo mais de 154 mil toneladas na última safra (CONAB, 2023). Esse crescimento deve-se aos avanços no melhoramento dessa cultura que resultaram em um número significativo de cultivares amplamente adaptados, altamente produtivos, com maior resistência a patógenos (EMBRAPA, 2015), e ao crescente investimento em tecnologia e pesquisa associado ao intercâmbio de informações técnicas entre empresas públicas e privadas, resultando em um aumento na eficiência do manejo da cultura (MEDINA *et al.*, 2016).

Para atender às diversas demandas do mercado, os programas de melhoramento de soja têm estabelecido diferentes objetivos, os quais incluem o desenvolvimento de cultivares transgênicas, com resistência a doenças e pragas, tolerância a estresses abióticos e melhoria na qualidade nutricional (SILVA *et al.*, 2017). Todos esses enfoques visam a obtenção de cultivares superiores, resultando em um aumento de produtividade média de 15,9% em relação a um aumento de área plantada de 6,2% na safra 2022/23 (CONAB, 2023). Neste sentido, é necessário aumentar a produção, de modo a acompanhar a grande demanda global considerando as estimativas de aumento populacional nos próximos anos (MASUDA E GOLDSMITH, 2009; RAY et al., 2013, SAGAE, 2022).

Uma das formas de aumentar a produção é através do melhoramento genético. Este é composto de várias etapas e processos, incluindo o desenvolvimento de populações, avaliações de linhagens e processos de seleção (ALMEIDA et al., 1999). Para que o programa de melhoramento obtenha sucesso, as etapas iniciam-se com a escolha dos parentais, cujo objetivo é montar um bloco de cruzamento com indivíduos superiores e, desta forma, obter uma população base de ampla variabilidade genética, com características de interesse e adaptabilidade (MUNIZ, 2007). Essa escolha deve ser baseada no desempenho “per-se” relativo às características de interesse,

capacidade combinatória, adaptabilidade e diversidade entre o grupo de genitores (SILVA et al., 2017).

Existem vários métodos de condução de populações segregantes, que podem ser divididos entre os que separam as fases de endogamia e seleção e os que não separam essas fases (BORÉM et al., 2017). Normalmente os métodos de condução variam entre os programas de melhoramento, podendo-se associar mais de um método de condução e variando a estratégia de avanço de gerações conforme disponibilidade de tempo, recurso financeiro, mão-de-obra e área para conduzir as populações.

A eficiência no processo de seleção está associada aos ganhos genéticos obtidos nos programas de melhoramento, através do cruzamento de genótipos superiores e divergentes, para as variáveis de interesse, explorando-se a variabilidade genética das populações (VOLPATO, 2016). Estudos como o de Felipe et. al (2016), mostram que o ganho genético médio anual nos principais países produtores de soja não acompanha a crescente demanda na produção de alimentos mundial.

Um passo de suma importância nos programas de melhoramento é a qualidade da fenotipagem dos genótipos, além de também encontrar características altamente correlacionadas com a produtividade para, assim, proporcionar novas estratégias de seleção que possam resultar em um maior ganho genético médio anual. Fiorani e Schurr (2013), referem-se à fenotipagem de plantas como o conjunto de protocolos e metodologias utilizados para medir, de forma acurada, o desenvolvimento, arquitetura e composição das plantas em diversas escalas.

Além de produtividade, outras variáveis devem ser consideradas para selecionar e/ou para caracterizar o material, como ciclo, hábito de crescimento, deiscência de vagem, quantidade de semente por vagem, resistência ao acamamento, cor de flor e de pubescência, altura de planta, altura da inserção da primeira vagem, formato da folha, dentre outras. Desta forma torna-se possível a recomendação de genótipos para ambientes específicos.

Segundo Cruz et al. (2012), a seleção pode ser direta, com base em características específicas, ou indireta, utilizando variáveis auxiliares. Contudo, se as correlações entre as variáveis auxiliares e a variável alvo forem de pequenas magnitudes, o uso de índices de seleção se torna primordial.

A utilização de índices de seleção foi inicialmente proposta por Smith (1936), quando ele observou que a seleção para caracteres quantitativos eram um problema

nos programas de melhoramento, devido a sua complexidade e dificuldade de determinar os valores genotípicos de plantas ou linhagens individuais, a menos que tivessem sementes e instalações suficientes para montar ensaios com repetições. O índice de seleção é um caráter adicional, estabelecido a partir da combinação ótima de vários atributos, que permitem realizar a seleção simultânea de variáveis múltiplas, de forma eficiente, obtendo-se ganhos equilibrados para as variáveis de importância econômica (CRUZ et al., 2012).

Ao longo do tempo, vários índices de seleção foram propostos, com o intuito de corrigir limitações dos índices já existentes. Atualmente alguns índices de seleção multicaracterísticos são baseados em ideótipos. Conceituado por Donald (1962) como “planta ideal” aquela que possui características específicas melhoradas para um ambiente ou conjunto de condições particulares. Tal conceito revolucionou o melhoramento de plantas ao mudar o foco da seleção de características individuais para a otimização da combinação de características de maneira a aprimorar o desempenho da cultura. Neste sentido, Olivoto e Nardino (2021), propuseram o índice MGIDI (Multitrait Genotype Ideotype Distance Index), o qual estima a distância genótipo-ideótipo, onde quanto menor a distância, mais próxima do ideótipo planejado. O valor máximo para as características de interesse é 100, e essa pontuação pode ser ajustada ponderadamente, levando em consideração a importância relativa de cada característica e os objetivos específicos da seleção.

A eficiência da seleção está intrinsecamente ligada à precisão da avaliação das variáveis de interesse. No entanto, como essa avaliação depende de mão de obra capacitada para obter alta acurácia, uma alternativa promissora para reduzir erros humanos nas avaliações, é a adoção da fenotipagem de alto rendimento (ARAUS & CAIRNS, 2014) com uso de câmeras de alta resolução acopladas a aeronaves não tripuladas. Esse tipo de fenotipagem vem sendo empregado em vários programas de melhoramento de soja (CASAGRANDE, 2019). Ela vem sendo utilizada para reduzir custos e agilizar o processo de avaliação dos genótipos dos programas de melhoramento, que podem possuir um alto custo e estar propício à subjetividade de avaliação e erros humanos (VOLPATO, 2020).

Segundo Araus et. al (2018), atualmente a maioria do que é considerado fenotipagem de alto rendimento, baseia-se em sensoriamento remoto. Esses autores enfatizam que os dispositivos mais utilizados para a fenotipagem de culturas

englobam câmeras que utilizam da radiação refletida ou emitida pelo dossel, sensores de fluorescência, térmicos, multiespectrais e hiperespectrais.

A escolha dos sensores desempenha um papel crucial no design de experimentos de fenotipagem e varia de acordo com os diversos objetivos dos experimentos. Cada tipo de sensor possui vantagens e desvantagens específicas. Por exemplo, as imagens RGB são econômicas e amplamente empregadas para medir biomassa, crescimento de plantas e características morfológicas (YANG et al., 2014).

A partir das imagens coletadas, são extraídas informações de interesse, como os índices de vegetação que podem ser utilizados como características secundárias na seleção dos programas de melhoramento, como quantidade de vegetação e condições fisiológicas da planta, para a seleção de características primárias, como por exemplo produtividade. Esta seleção indireta é baseada em características específicas, que compõem características agrônômicas relevantes como por exemplo a utilização de NDVI, índice de vegetação de diferença normalizada, para a avaliação de biomassa; utilização de CI, índice de clorofila, para avaliar conteúdo de clorofila (LI & SHI., 2018) dentre outras possibilidades.

Esses índices não apenas servem para avaliar características das plantas, mas também podem ser empregados na identificação de alvos de interesse, como solo exposto, água e a própria vegetação, como é o caso do índice GLI (Green Leaf Index), um índice de folha verde. O índice GLI foi posteriormente empregado por Volpato (2021) para modelar o decréscimo de verde e avaliar o ciclo de genótipos de soja.

Neste sentido os objetivos deste trabalho foram: definir os ideótipos de soja de diferentes grupos de maturação ao incluir o estudo de características agrônômicas e de curvas de maturação em progênies de soja e comparar diferentes estratégias de seleção utilizando as características agrônômicas e da curva de maturação, para selecionar genótipos superiores em cada grupo de maturação.

## **2. MATERIAL E MÉTODOS**

### **2.1. Material vegetal, descrição do local e coleta de dados**

Durante a safra 2022/23 foram realizados dois ensaios de teste de progênies F<sub>3:5</sub> do Programa de Melhoramento de Soja da Universidade Federal de Viçosa, as quais foram selecionadas devido a produtividade e precocidade apresentadas. Os ensaios foram conduzidos nas Unidades de Ensino, Pesquisa e Extensão (UEPE) da

Universidade Federal de Viçosa, denominadas UEPE Vale da Agronomia (população 1 - VA) (20° 46'08" S, 42° 52' 13' W) e na UEPE Horta Nova (população 2 - HN) (20° 45' 46" S, 42° 49' 19" W), ambas no município de Viçosa-MG (Figura 1).

Os experimentos seguiram o delineamento de blocos aumentados (DBA), com parcelas compostas de 3 linhas de 1,5 m de comprimento, espaçamento de 0,5 m entre linhas e densidade populacional de 16 plantas/linha.

A população 1 foi constituída de 500 famílias e cinco cultivares comerciais utilizadas como testemunha (BRASMAX Cromo TF IPRO, M5838 IPRO, NEO610 IPRO, M6410 IPRO e DM 66I68 IPRO), distribuídas em 550 parcelas e 10 blocos. A semeadura foi realizada no dia 28/10/2022. As 500 famílias foram originadas de 84 populações provenientes de 8 cultivares comerciais e 7 linhagens do Banco de Germoplasma do Programa de Melhoramento da Soja do Departamento de Agronomia da Universidade Federal de Viçosa, utilizadas como genitores.

A população 2 foi constituída de 264 famílias e seis cultivares comerciais que foram utilizadas como testemunha (BRASMAX Cromo TF IPRO, Apolo RR, M5838 IPRO, NS5909, TMG 7262RR e TMG 7363 RR) distribuídas em 300 parcelas e 6 blocos. A semeadura foi realizada no dia 23/11/2022. As 264 famílias foram originadas de 15 populações, obtidas por meio de um dialelo completo em que foram utilizadas seis cultivares comerciais como genitores.

Durante a condução dos ensaios foram avaliados os seguintes caracteres agronômicos: altura e data de florescimento, acamamento de plantas, altura e data de maturação, peso de 100 sementes e peso total de grãos. Os caracteres foram aferidos como se seguem:

a) altura de planta na maturação (ALTMAT) em centímetros, utilizou-se uma régua de campo e foi aferida a altura a partir do solo até o final da haste principal;

b) número de dias para florescimento (NDF), calculada da data do plantio de cada ensaio até a data de florescimento, considerando pelo menos 50% das da parcela plantas com flores;

c) acamamento das plantas (ACAM), avaliado utilizando-se uma escala variando de 1 a 9, onde 1 representava as plantas suscetíveis e 9 as plantas resistentes ao acamamento;

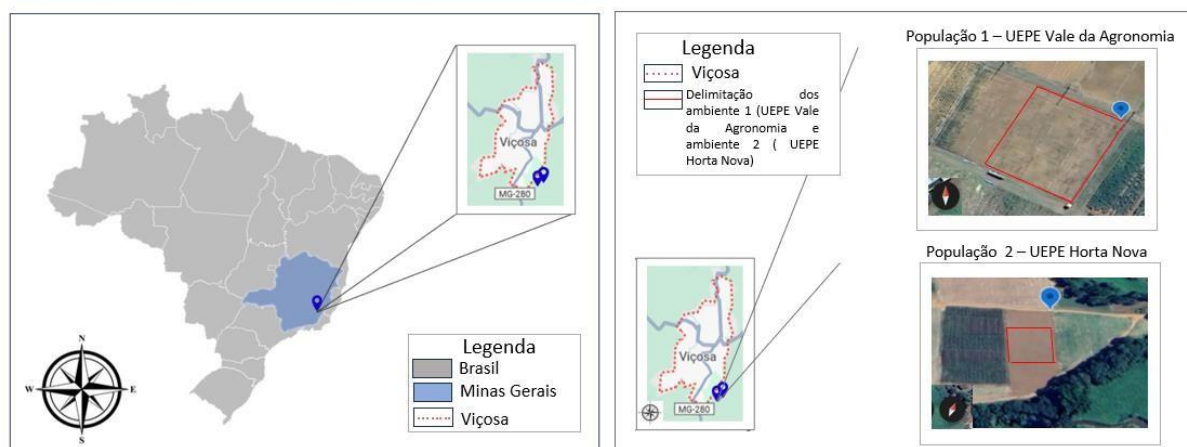
d) número de dias para maturação (NDM), calculada da data do plantio de cada ensaio até a maturação fisiológica da parcela, considerando pelo menos 50% das plantas no estágio de desenvolvimento R8, segundo Fehr e Caviness (1977);

e) peso de 100 sementes em gramas (P100), em que 100 sementes foram pesadas em balança analítica;

f) peso total de grãos em gramas (PROD), em que os grãos produzidos em cada parcela foram pesados individualmente em balança analítica.

Para os caracteres P100 e PROD utilizou-se a padronização da umidade dos grãos em 13%, aferindo-se a umidade com o Medidor de umidade de grãos John Deere GT5300.

Os ensaios foram monitorados semanalmente para a detecção e controle de pragas e doenças. Todas as práticas de manejo foram realizadas seguindo as recomendações para a cultura da soja (SILVA, et al., 2022), conforme necessidade de cada área.



**Figura 1.** Município de Viçosa-MG, e População 1 (UEPE Vale da Agronomia) e População 2 (UEPE Horta Nova). Fonte: Google Earth e Google Maps.

## 2.2. Planejamento, execução dos voos e avaliação da precisão do georreferenciamento

O aerolevanteamento de ambos os ambientes foi realizado utilizando-se o VANT (Veículo Aéreo Não Tripulado) DJI Phantom 3 Profissional (DJI Technology Co., Ltd., Shenzhen, China), equipado com a câmera RGB com resolução de 12,4 megapixels e lente de 20 mm com abertura óptica f/2.8.

Em cada experimento foi feito um plano de voo utilizando o software de planejamento de voo Pix4Dcapture (v4.8.0; Pix4D, Prilly, Suíça), onde os pontos de passagem e as rotas de voo foram gerados automaticamente. Desta forma, durante cada voo, o VANT percorreu uma rota predeterminada. Todos os voos foram realizados a uma altura de 20 m e na velocidade de 1,8 m/s. As imagens foram capturadas com uma sobreposição de 85% em todas as direções. Como resultado, as

imagens apresentavam uma distância de amostragem do solo (GSD) de 0,5 cm por pixel.

O planejamento de cada missão sempre levou em consideração a segurança de quem estava na área experimental durante os voos, a legislação vigente e as condições ambientais. Antes de cada voo, verificava-se a atividade solar (Índice K) utilizando o software Geomagnetic Storms (v1.43 APK; Caffeine Software) para garantir a segurança da operação. O limite seguro de voo era estabelecido em  $K=4$ , pois acima desse valor o GPS do VANT poderia deixar de funcionar corretamente e o veículo entraria em modo de altitude, onde o drone passa a ser controlado pelo piloto.

Os voos foram preferencialmente realizados entre 11:00 e 13:00 horas, pois as condições de iluminação são mais favoráveis e evita-se a interferência de sombras. Não foram realizados voos em dias chuvosos por questão de segurança e porque as gotículas de água sobre a folha interferem na absorção e reflectância da folha, podendo causar erro quando fossem extraídos os índices de vegetação.

Em cada experimento foram instalados Ground Control Points (GCPs). Esta é uma ferramenta fundamental no mapeamento com VANT para melhorar a precisão espacial, corrigir distorções e garantir a integração dos dados capturados pelo VANT com outras fontes de informação geográficas. Os GCPs foram estrategicamente distribuídos nas áreas dos experimentos, sendo marcados por alvos compostos por quadrados de 0,30x0,30m com um padrão xadrez preto e branco. O dimensionamento dos pontos de controle foi satisfatório visto à altura de voo utilizada. Esses alvos foram construídos com lona front light para garantir uma maior visibilidade e precisão durante o processo de captura das imagens. O experimento 1 possuía 7 GCP's e o experimento 2 possuía 8 GCP's.

As coordenadas geográficas de cada GCP's foram coletadas com o receptor GNSS RTK Topomap TPS T10 (experimento 1) e GNSS no método estático, onde o receptor ficou meia hora em cada ponto (experimento 2), utilizando-se o receptor TRIUMPH-1 JAVAD. As coordenadas de ambos os ensaios foram processadas pelo Processamento de Ponto Preciso (PPP) IBGE. As coordenadas foram coletadas no Datum SIRGAS 2000.

A frequência de voos e quantidade variou em cada uma das populações, totalizando 15 voos na população 1, e 18 voos na população 2. Os voos utilizados nas análises foram selecionados conforme a qualidade do ortomosaico gerado, sendo eliminados os que possuíam eventuais manchas de sombras.

**Tabela 1.** Relação da frequência de voos utilizados nas análises das Populações 1 (UEPE Vale da Agronomia) e 2 (UEPE Horta Nova) conduzidos durante a safra 2022/23.

UEPE Vale da Agronomia – População 1			UEPE Vale Horta Nova – População 2		
Nº do Voo	Data do Voo	DAP*	Nº do Voo	Data do Voo	DAP
1	28/10/22	0	1	02/12/22	9
2	09/12/22	42	2	11/12/22	18
3	12/12/22	45	3	11/01/23	49
4	21/12/22	54	5	19/01/23	57
5	30/12/22	63	5	24/01/23	62
6	12/01/23	76	6	01/02/23	70
7	24/01/23	88	7	14/02/23	83
8	26/01/23	90	8	07/03/23	104
9	02/02/23	97	9	09/03/23	106
10	22/02/23	117	10	13/03/23	110
11	24/02/23	119	11	21/03/23	118
12	28/02/23	123			

\*DAP: Dias após o plantio.

### 2.3 Processamento e análise de imagem

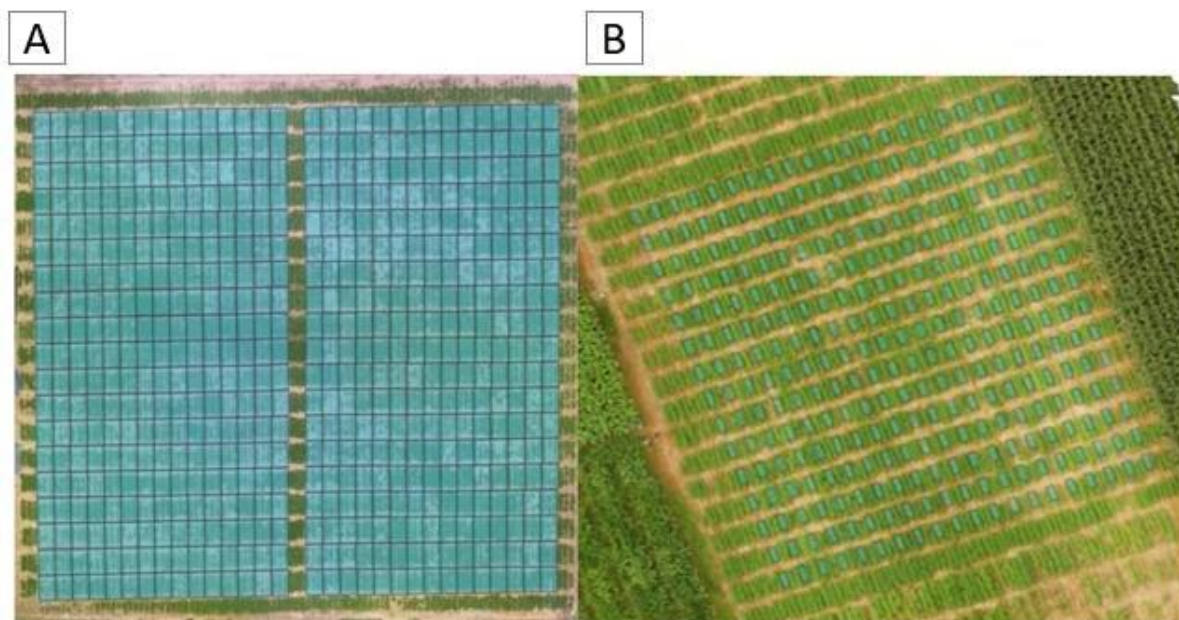
O processamento das imagens foi realizado no Agisoft Metashape Professional 64bit (Agisoft LLC, St. Petersburg, Rússia), envolvendo as etapas de alinhamento das imagens, geração da nuvem densa de pontos em alta qualidade, formação do mesh a partir da nuvem densa de pontos e criação do ortomosaico.

Na construção do ortomosaico, os pontos de controle foram adicionados à nuvem densa de pontos em baixa qualidade. Então as fotos foram ajustadas e alinhadas novamente. Foi feita uma nova nuvem densa de pontos em alta qualidade, e realizado o mesh que consiste em uma malha triangular que representa uma superfície tridimensional usando triângulos. Assim, o ortomosaico é construído com as coordenadas corrigidas.

Para o processamento foi utilizado o Datum WGS 84 zona 23 S, pois é o Datum internacional.

A grade shapefile, da população 1 foi construída no software QGIS (version 3.30, 2023). Cada célula da grade retangular foi desenhada sobre uma parcela experimental (2,00 m x 1,30 m). A grade shapefile da população 2 foi construída no software *R* utilizando o pacote FieldImageR extra. Para este último optou-se por utilizar apenas a linha central para extrair informações (Figura 2). Para a construção da grade, foi necessário montar um arquivo indicando o posicionamento de cada parcela no campo (bloco, linha, coluna e faixa). Além disso, o programa foi informado sobre o número total de linhas e colunas. Para os dados do ambiente 2 foi utilizado um buffer de -0.5, desta forma permanecendo apenas a linha central da parcela,

evitando ter influências das parcelas vizinhas. Adicionalmente, se atentou a colocar o valor máximo de pixels igual a 118672044, devido ao tamanho do ortomosaico.



**Figura 2.** Grade Shapefile: a) ensaio localizado na Uepe Vale da Agronomia (população 1); b) ensaio localizado na Uepe Horta Nova (população 2).

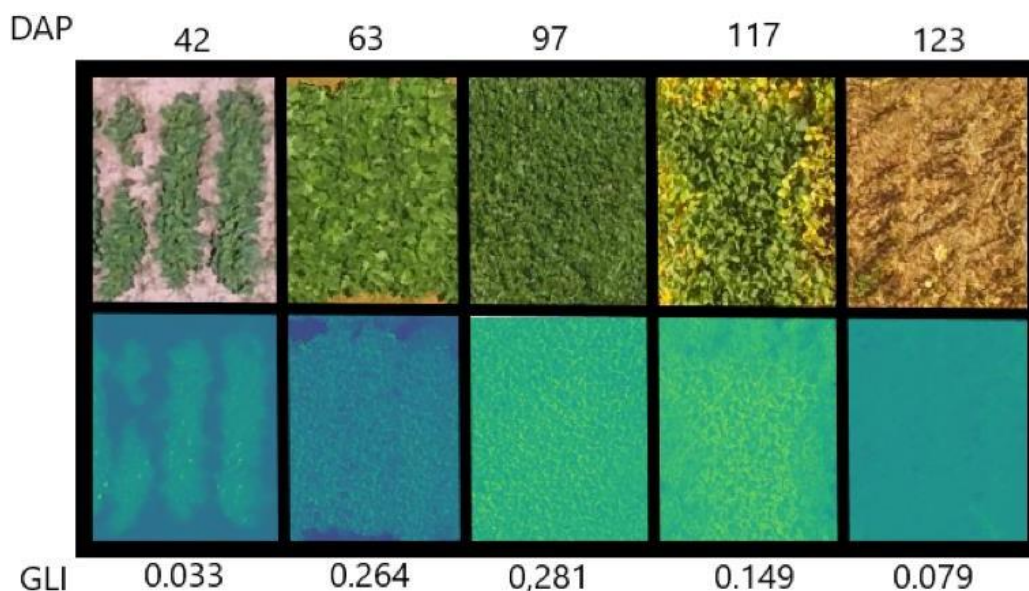
Todos os ortomosaicos foram analisados no software R. Foram utilizados os pacotes FIELDimageR (MATIAS et. al, 2020), FIELDimageR.Extra (PAWAR e MATIAS, 2023), terra (HIJIMANS, 2020), mapview (APPELHANS et.al, 2016), sf (PEBESMA, 2018) e stats (PEBESMA,2023) para a manipulação dos dados em raster e do polígono (shapefile).

De cada imagem (parcela) foram extraídos os valores de R, G, B e posteriormente, e a partir destes valores, foram calculados os valores médios do índice GLI de cada parcela. Para isso foi preciso especificar a sequência das bandas vermelho, verde e azul; ou 1, 2 e 3. O processo de obtenção dos dados foi executado nos voos selecionados, sendo realizado separadamente para cada parcela a fim de construir a curva de maturação (Figura 3).

O índice GLI foi proposto por Louhaichi et al. (2001), na identificação de áreas de pastejo de gansos, com objetivo de separar planta de solo. Segundo Louhaichi et al (2001), se o valor do GLI for positivo ele tende a indicar presença de folhas verdes ou caules. Se o valor for negativo representa solo ou planta sem vida. Em grande parte dos casos pode-se utilizar uma linha de corte próximo a zero para separar as

duas classes. Por essa razão este índice pode ser utilizado em estudos de mudança na cobertura vegetal.

$$GLI = \frac{2 * G - R - B}{2 * G + R + B}$$



**Figura 3.** Imagem real e imagem em escala de falsa cor após extração do índice de vegetação GLI em uma parcela experimental, representando o desenvolvimento da cultura em cinco períodos correspondentes a diferentes dias após o plantio (DAP), safra 2022/23, Viçosa-MG.

### 2.3.1 Obtenção de dados baseados na curva de maturação construída por meio do índice GLI

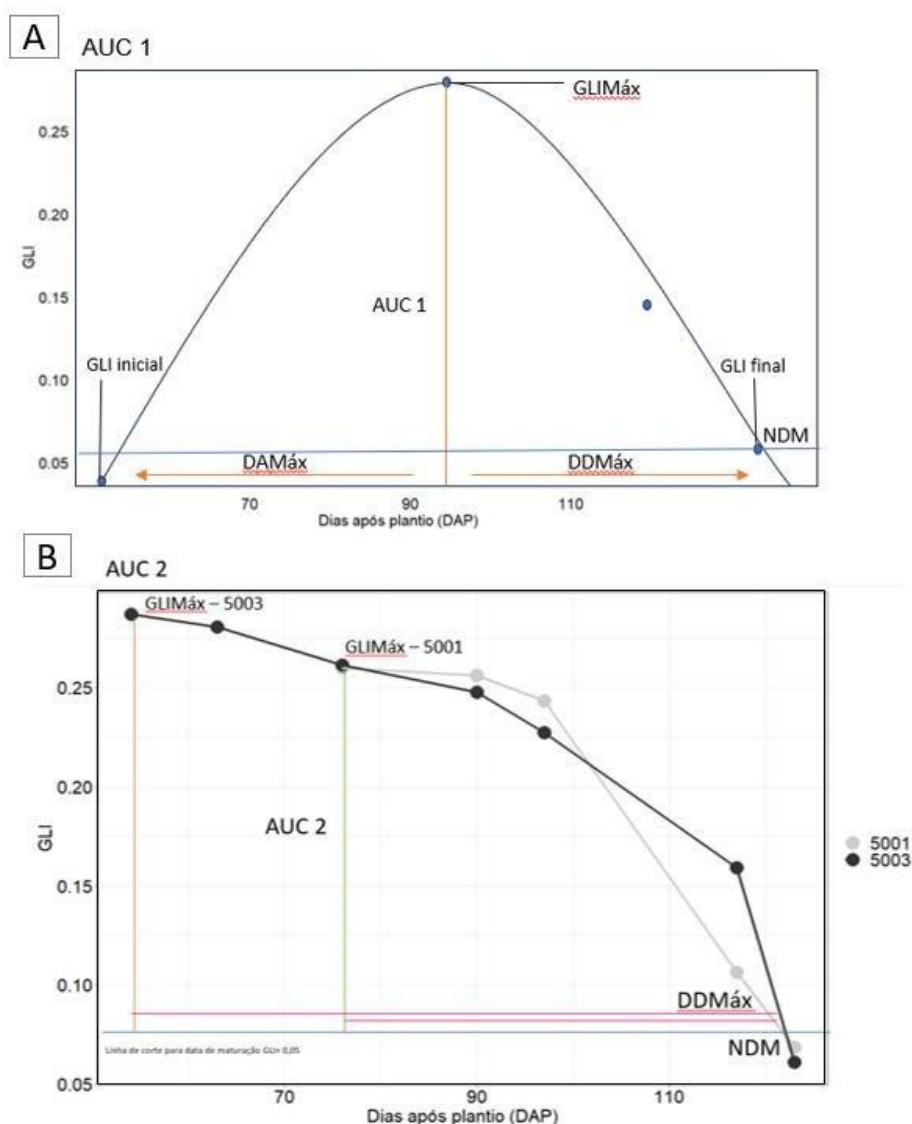
Para traçar a curva de maturação foram utilizados os valores da reflectância do índice GLI em função dos dias dos voos, computados como dias após o plantio (DAP). Os pontos foram ajustados através da regressão *LOESS* (Regressão Local), método não paramétrico que estima curvas através da suavização local (CLEVELAND, 1979).

Após a aplicação do método de suavização, para cada genótipo foram obtidas as seguintes características: área total abaixo da curva (AUC 1), o ponto de máximo GLI (GLIMáx), o número de dias até o ponto de máximo GLI (DAMáx), o número de dias do ponto de máximo GLI até a maturação (DDMáx) e a área embaixo da curva, obtida a partir do ponto de máximo GLI até a maturação completa (AUC 2), e verde acumulado, calculado pela diferença entre o GLIMáx e o valor inicial do índice GLI (Figura 4A).

A AUC 1 foi utilizada como característica para discriminar e observar as diferentes trajetórias de crescimento dos genótipos, pois alguns possuem melhor desempenho nas fases iniciais do desenvolvimento e podem reduzir a competitividade ao longo do ciclo, enquanto outros não possuem bom desempenho nas fases iniciais, mas vão adquirindo competitividade ao longo do ciclo. A AUC 2, foi utilizada para verificar a velocidade da senescência de cada genótipo. Quanto menor a AUC 2, dentro do grupo de genótipos de mesmo ciclo, mais rápida a senescência do genótipo, conforme representado para dois genótipos na Figura 4B.

Ambas as áreas abaixo da curva foram calculadas utilizando a função “Integrate” no software R, pacote stats (Stats: The R Stats Package, 2023), a qual pode ser utilizada para estimar a área sob uma curva de uma função específica em um intervalo definido. Neste sentido, foi ajustado o modelo LOESS para cada parcela, utilizando-se os dados de GLI e DAP, calculado os valores preditos da variável dependente (GLI) com base nos valores da variável independente (DAP) e criado uma função linear aproximada com a função “approxfun”, do pacote stats, baseada nos valores previstos de GLI usando o modelo Loess ajustado, para uma sequência de valores de DAP, com incremento de 0.01.

A partir desta curva foram calculadas as AUC 1 e AUC 2. Os limites inferior e superior estabelecidos na integração para o cálculo da AUC 1 foram os valores mínimos e máximos de DAP de todos os voos. Para cada parcela para a AUC 2 os limites inferiores foram o DAP correspondente ao GLIMáx de cada parcela e o limite superior foi o valor máximo do DAP. Este valor máximo de DAP corresponde aos NDM de cada parcela, desta forma variando conforme o ciclo de cada genótipo.

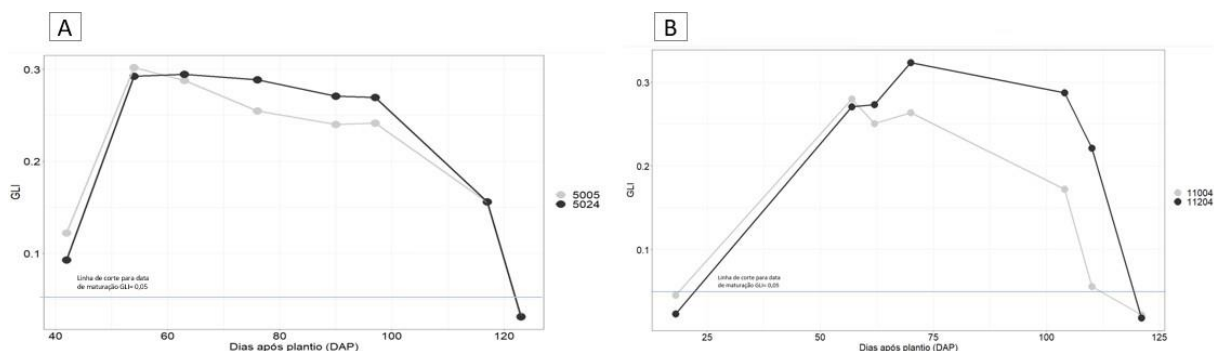


**Figura 4.** Representação gráfica das AUC 1, AUC 2 e variáveis da curva de maturação. A: Área total abaixo da curva (AUC 1), valor máximo para o índice de vegetação GLI (GLIMáx), número de dias até o valor máximo do índice de vegetação GLI (DAMáx), número de dias depois do valor máximo do índice de vegetação GLI (DDMáx), número de dias para maturação (NDM) na curva ajustada pelo modelo de regressão LOESS. B: área em baixo da curva a partir do valor máximo do índice de vegetação GLI (AUC 2), antes do ajuste pelo modelo de regressão LOESS, para os genótipos 5001 e 5003, com valores de máximo atingidos em 76 e 54 dias após plantio (DAP) respectivamente.

## 2.4 Análises estatísticas

Visando a melhor compreensão da influência das características agronômicas e das obtidas por meio do índice de vegetação GLI, os genótipos foram agrupados dentro de cada ensaio conforme avaliação de data de maturação realizada em campo. Essa abordagem foi adotada para assegurar uma análise justa, mitigando potenciais

distorções que poderiam ocorrer da comparação entre genótipos de diferentes grupos de maturação (Figura 5).



**Figura 5.** Representação gráfica da razão dos genótipos serem comparados e selecionados dentro de cada grupo de data de maturação estimada. Em A, genótipos com mesma data de maturação; em B, um genótipo mais tardio (11204) e um mais precoce (11004).

Após a formação dos grupos, todas as análises subsequentes foram conduzidas individualmente para cada grupo. Devido à não conformidade dos dados com as premissas da ANOVA, de normalidade dos dados e homogeneidade das variâncias, optou-se por utilizar o teste de Kruskal-Wallis (1952). Quando os resultados do teste de Kruskal-Wallis foram estatisticamente significativos, foram realizados testes de comparações múltiplas de Dunn (1964), comparando cada par de grupos entre si, com o objetivo de identificar quais grupos apresentavam diferenças estatisticamente significativas, com um nível de significância estabelecido em  $\alpha=5\%$ .

#### 2.4.1 Correlação entre as variáveis

As variáveis extraídas da curva de maturação e as variáveis agrônomicas foram correlacionadas entre si, pelo método da correlação linear de Pearson ao nível de 5% de probabilidade pelo teste F.

O coeficiente de correlação foi obtido pela equação:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum (x_i - \bar{x})^2)(\sum (y_i - \bar{y})^2)}}$$

em que:

$r$  = coeficiente de correlação de Pearson;

$x$  e  $y$  são os valores das duas variáveis para cada observação  $i$  ( $i = 1, \dots, n$ );

$\bar{x}$  e  $\bar{y}$  representam a média dos valores das variáveis  $x_i$  e  $y_i$ , respectivamente.

### 2.4.2 Seleção dos genótipos

A seleção dos genótipos foi realizada dentro de cada grupo e foi realizada utilizando-se duas etapas. Na primeira etapa, foi realizado o ajuste de um modelo misto para obtenção dos BLUPS (*Best Linear Unbiased Prediction*) das progênies utilizando o pacote lme4 (BATES et al., 2015). No modelo as testemunhas foram consideradas como fixas e os efeitos de bloco e progênies como aleatórios. Na segunda etapa, utilizou-se o pacote Metan (OLIVOTO, 2020) para seleção das progênies utilizando o índice MGIDI (OLIVOTO & NARDINO, 2020), considerando as variáveis da curva de maturação com sentidos de seleção variando entre os grupos.

Segundo Olivoto & Nardino (2020), a aplicação do índice MGIDI baseia-se em quatro etapas: i) reescalonamento das unidades: todas as características são redimensionadas para uma escala de 0 a 100, para que possam ser avaliadas em conjunto; ii) análise fatorial, utilizada para explicar a estrutura de correlação e redução da dimensionalidade dos dados; iii) estabelecer um ideótipo com base no valor de dados conhecido /desejado de cada característica; e, iv) calcular a distância entre o genótipo e o ideótipo pré estabelecido na etapa anterior, a partir da seguinte equação:

$$MGIDI_i = \sqrt{\sum_{j=1}^f (F_{ij} - F_j)^2}$$

Onde: MGIDI<sub>i</sub>: Índice MGIDI no i-ésimo genótipo;

F<sub>ij</sub>: pontuação do i-ésimo genótipo no j-ésimo fator (i = 1, 2, ..., g; j = 1, 2, ..., f), sendo g e f o número de genótipos e fatores, respectivamente;

F<sub>j</sub>: o j-ésimo pontuação do ideótipo.

O genótipo com menor MGIDI é o mais próximo do ideótipo e, conseqüentemente, apresentará os valores desejados para todas as características analisadas

A partir da distância genótipo-ideótipo foi feito o ranqueamento dos melhores genótipos para cada grupo. Foram realizadas sete seleções com diferentes critérios (Tabela 2).

**Tabela 2.** Relação das seleções e os critérios de seleção utilizados n (UEPE Vale da Agronomia e UEPE Horta Nova).

Seleção	Critério
1	Seleção direta através dos BLUPS para produção (PROD) de cada progênie
2	Seleção pelo índice MGIDI, utilizando-se apenas as variáveis da curva de maturação
3	Seleção pelo índice MGIDI, utilizando-se apenas as variáveis agronômicas, exceto produção (PROD)
4	Seleção pelo índice MGIDI, utilizando-se as variáveis da curva de maturação e as variáveis agronômicas, exceto produção (PROD)
5	Seleção pelo índice MGIDI, utilizando-se as variáveis da curva de maturação e produção (PROD)
6	Seleção pelo índice MGIDI, utilizando-se as variáveis agronômicas e a produção (PROD)
7	Seleção pelo índice MGIDI, utilizando-se as variáveis da curva de maturação e as variáveis agronômicas acrescidas da produção (PROD)

Em cada seleção, foi verificado a porcentagem de coincidência do ranqueamentos dos dez melhores genótipos das seleções 2 a 7 em relação à seleção 1. Este procedimento proporcionou uma análise abrangente da eficácia de diferentes critérios de seleção na identificação dos genótipos mais promissores. As estratégias de seleção foram comparadas através do ganho genético obtido em peso de parcela.

O ganho genético foi estimado pela equação:

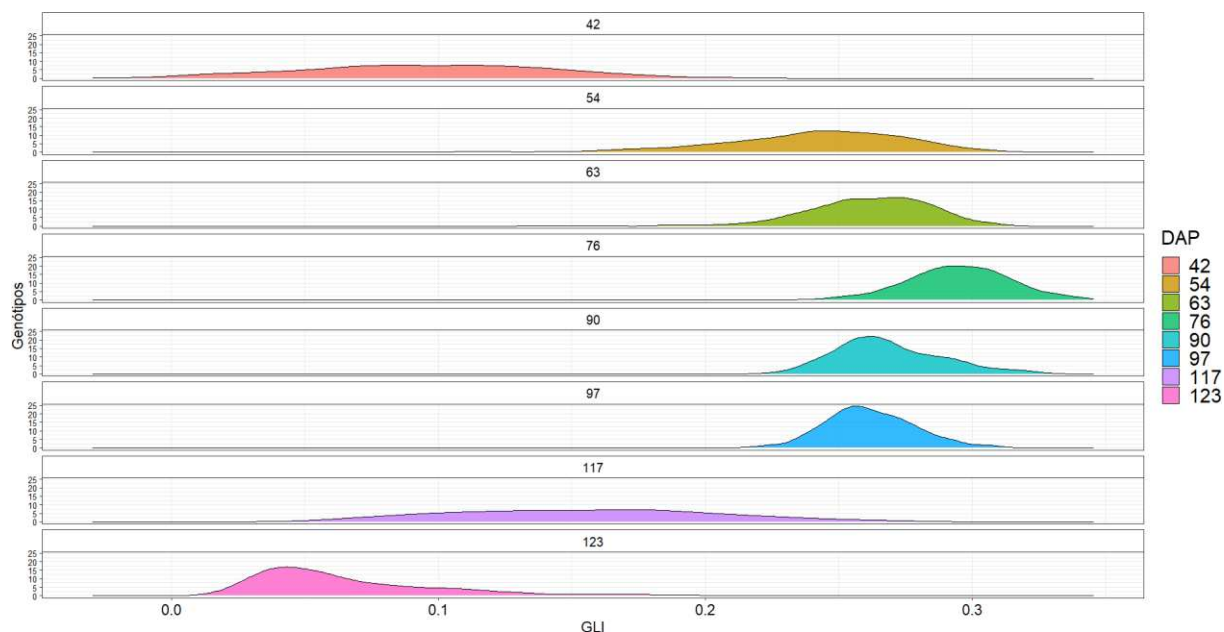
$$GS = h^2 \times DS$$

em que, DS é o diferencial de seleção, obtido pela diferença da média da população inicial e a média das dez melhores progênies de cada seleção; e  $h^2$  é a herdabilidade da variável, que representa a proporção da variação fenotípica que é devido à variação genotípica.

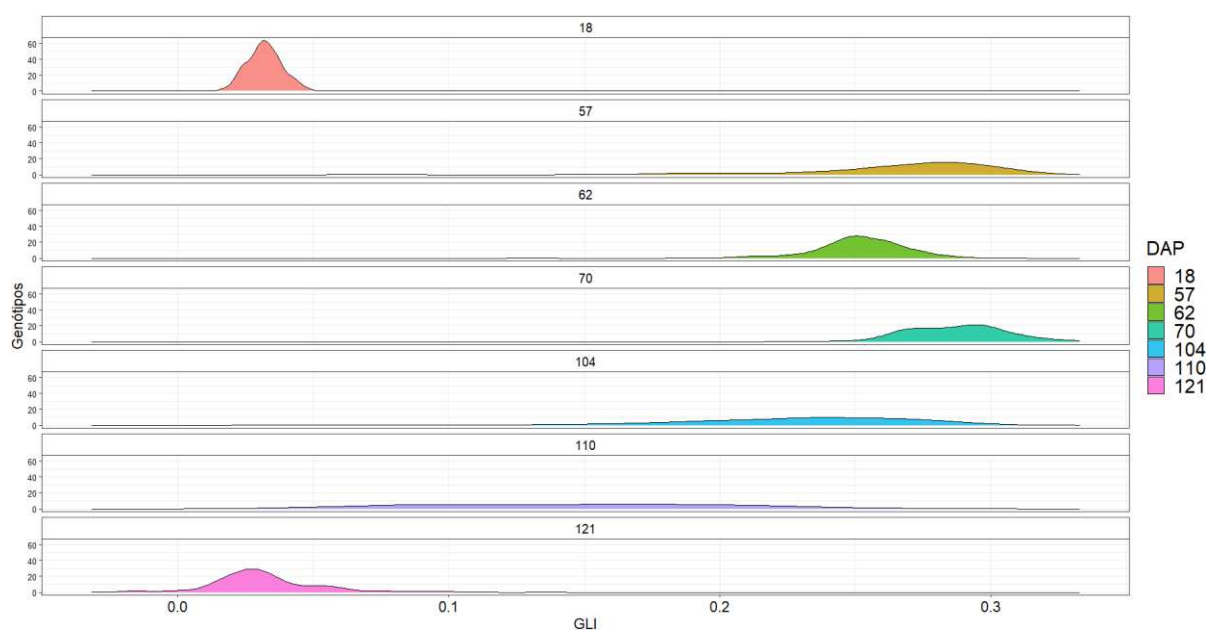
### 3. RESULTADOS E DISCUSSÃO

#### 3.1 Variáveis da curva de maturação

Nas Figuras 6 e 7 são apresentadas as distribuições dos genótipos, baseadas nos valores médios do índice GLI de todos os genótipos ao longo dos voos, para as populações 1 e 2, mostrando que o índice GLI foi eficiente para modelar o decréscimo do verde, assim como o estudo de Volpato et al. (2021), que utilizou câmera RGB, e o índice GLI para modelagem do decréscimo de verde para estimar data de maturação em soja.



**Figura 6.** Distribuição dos genótipos baseado nos valores de refletância do índice GLI ao longo do ciclo na População 1. DAP: Dias após maturação.



**Figura 7.** Distribuição dos genótipos baseado nos valores de refletância do índice GLI ao longo do ciclo na população 2. DAP: dias após maturação.

De modo geral as progênies avaliadas da população 1 possuem ciclo mais longo que as progênies avaliadas na população 2. Essa observação pode ser observada pelo maior número de genótipos com valor acima de 0.2 do índice GLI aos 117 DAP.

Na população 1, pode-se observar que aos 42 DAP, os genótipos já estavam distribuídos entre valores próximos a zero e 0.2 de GLI, indicando que as plantas estavam se desenvolvendo e que alguns genótipos já estavam fechando dossel, reduzindo-se a porção de solo exposto e aumentando o valor do índice.

A maior concentração dos maiores valores de GLI foi em 76 DAP, ficando em torno de 0,28 e após essa data observamos que os genótipos foram entrando em senescência até chegar em valores mínimos em 123 DAP. Na população 2, para os voos de 18 DAP e 121 DAP, observa-se que os valores de GLI ficaram abaixo de 0,05 o que indica muita exposição do solo e que as plantas já estavam em R8, respectivamente.

Para os voos intermediários observa-se uma variação nos valores do índice para os genótipos, indicando que as plantas estavam em crescimento até chegar no seu valor máximo que iniciou a partir de 49 DAP e se estendeu até os 110 DAP, diferente da população 1.

O comportamento do índice se mostrou similar ao encontrado por Louhaichi (2001), quando foi utilizado para identificar áreas de pastejo de gansos, se mostrando um índice eficiente para separar planta de solo.

O número de dias para a maturação foi utilizado para a formação dos grupos de genótipos para cada ensaio avaliado. Na população 1 foram formados quatro grupos, sendo o grupo 1 o mais precoce e 4 o mais tardio. A variação entre dias de maturação entre o genótipo mais precoce do grupo 1 e o genótipo mais tardio do grupo 4 foram de 20 dias. Segundo a Embrapa (2008), os grupos de maturação de soja para Minas Gerais podem ser classificados como semiprecoce (101 a 110 dias); médio (111 a 125 dias); semitardio (125 a 145 dias) e tardio (> 145 dias). Neste sentido os genótipos do grupo 1 foram classificados como semiprecoces, e os dos grupos 2, 3 e 4 classificados como de ciclo médio.

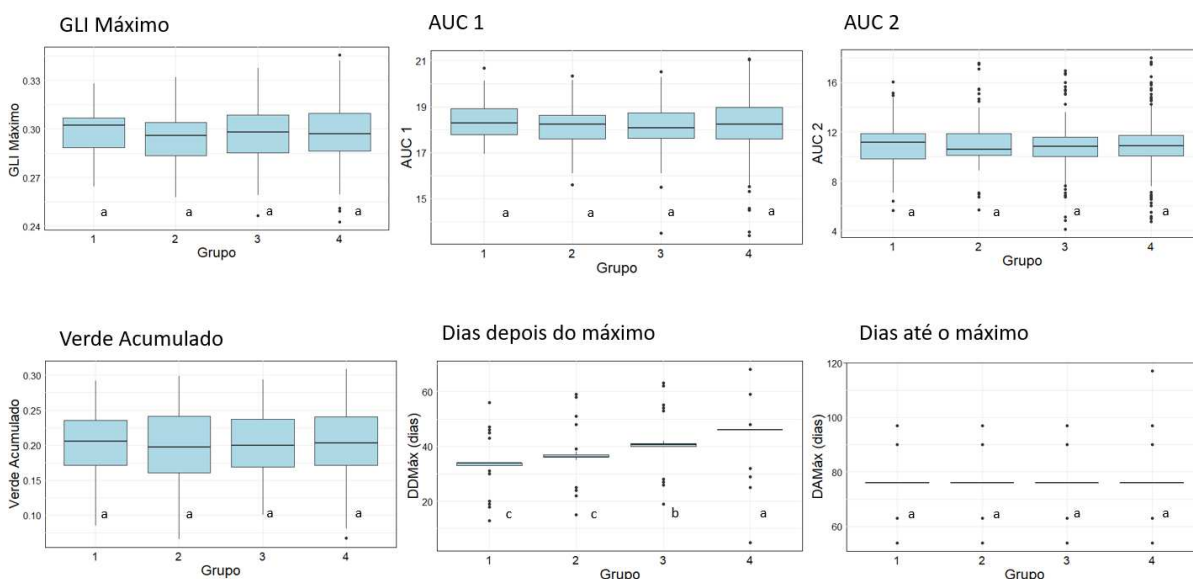
A população 2 seguiu a mesma metodologia para a formação de grupos. Onde, também foram formados 4 grupos, com variação de 19 dias entre o genótipo mais precoce do grupo 1 e o mais tardio do grupo 4. Segundo a classificação da Embrapa (2008), o grupo 1 possui genótipos semiprecoces e os grupos 2, 3 e 4 possui genótipos de ciclo médio.

Nas análises do realizadas de teste de Dunn e gráficos de boxplot, das variáveis da curva de maturação para as populações 1 e 2 (Figuras 8 e 9) observa-se que na população 1, DDMáx foi a única variável que apresentou diferença estatística

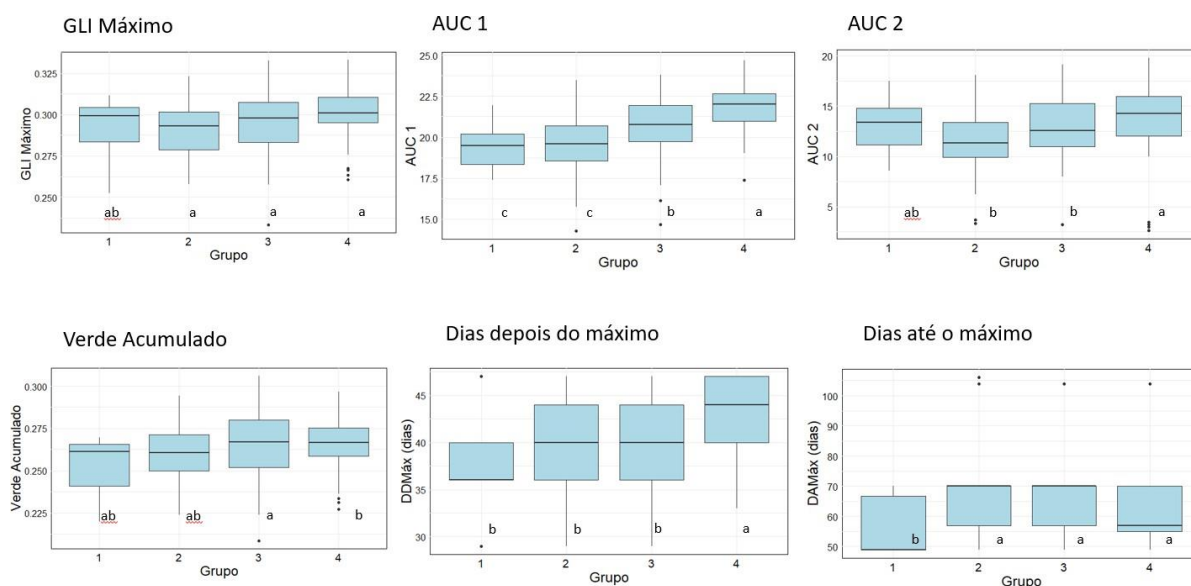
entre as médias dos grupos, onde o grupo 4 apresentou a maior média, seguido do grupo 3, e os grupos 2 e 1 tenderam a apresentar valores de média igual. O grupo 4 foi o que obteve maior quantidade de outliers para todas as variáveis, e todos os grupos apresentaram outliers para AUC 2.

Para verde acumulado, o grupo 2 foi o que apresentou maior amplitude entre quartil, demonstrando maior variação em comparação aos demais grupos. DAMáx apresentou o mesmo comportamento para os quatro grupos. Já a variável DDMáx se comportou conforme o ciclo dos grupos, onde o grupo 1 (mais precoce) apresentou a menor média, e o grupo 4 (mais tardio) apresentou maior média.

Na população 2, todas as variáveis apresentaram diferença entre as médias dos grupos. Para AUC 1, o grupo 4 foi superior aos grupos 3, 2 e 1, seguido do grupo 3, e os grupos 1 e 2 não apresentaram diferença estatística entre as médias segundo o teste de Dunn. Para AUC 2, a média do grupo 4 foi equivalente ao grupo 1, e superior aos grupos 2 e 3, mas o grupo 1 foi equivalente aos grupos 2 e 3. Para a variável DAMáx, os grupos 2, 3 e 4 tenderam a apresentar valores de média igual, já o grupo 1 foi inferior aos demais, o oposto ocorreu para a variável DDMáx, onde o grupo 1 foi superior aos demais grupos.



**Figura 8.** Boxplot do valor máximo para o índice de vegetação GLI (GLIMáx), Área total abaixo da curva (AUC 1), área em baixo da curva a partir do valor máximo do índice de vegetação GLI (AUC 2), verde acumulado, número de dias até o valor máximo do índice de vegetação GLI (DAMáx) e número de dias depois do valor máximo do índice de vegetação GLI (DDMáx) para os grupos 1 a 4, população 1.



**Figura 9.** Boxplot do valor máximo para o índice de vegetação GLI (GLIMáx), Área total abaixo da curva (AUC 1), área em baixo da curva a partir do valor máximo do índice de vegetação GLI (AUC 2), verde acumulado, número de dias até o valor máximo do índice de vegetação GLI (DAMax) e número de dias depois do valor máximo do índice de vegetação GLI (DDMax) para os grupos 1 a 4, população 2.

### 3.2 Variáveis agrônômicas

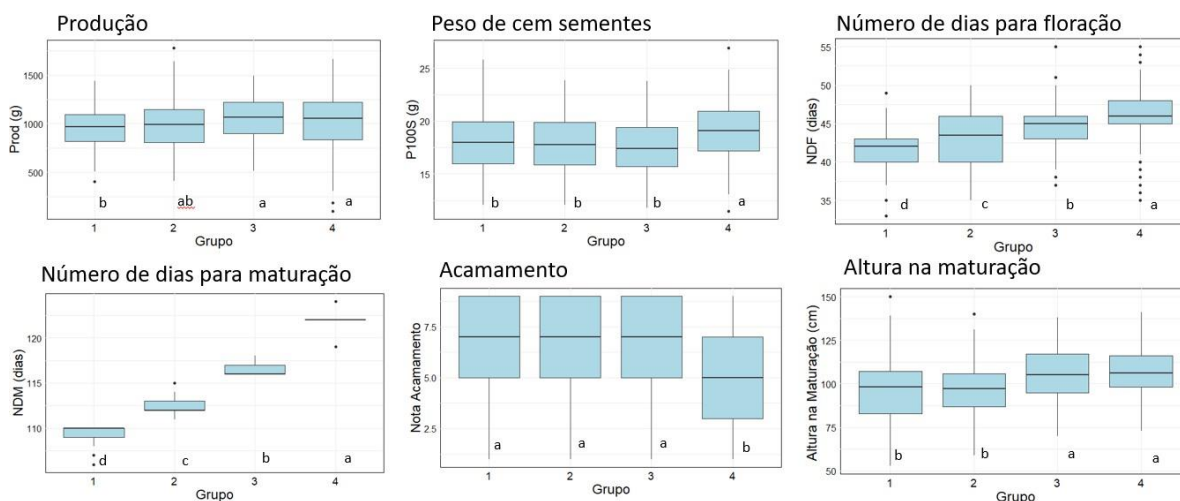
Na população 1 as variáveis produção e peso de 100 sementes não apresentaram variação expressiva entre os grupos, demonstrando que todos os grupos formados possuem genótipos superiores para serem selecionados visto que as médias variaram de 951,6 g para 1032,47 g em produção e 17,42 g para 18,88 g em P100 (Figura 10).

O grupo 1 foi o que apresentou genótipos com menores médias, segundo o teste de Dunn, para a variável NDM, já o grupo 4 as maiores médias. Devido a separação dos grupos pela data de maturação pode-se observar uma amplitude entre quartil baixa em todos os grupos.

O grupo 2 apresentou a maior variação para início da floração (NDF), tendo em vista a maior amplitude entre quartil apresentada no boxplot, apesar de os grupos não apresentarem diferença estatística para as médias desta variável.

O acamamento foi a característica que apresentou maior amplitude entre quartil, demonstrando que os grupos apresentaram todas as notas de acamamento (1 a 9). Contudo, pelo teste de Dunn, os grupos 1,2 e 3 tenderam a apresentar mesma média para esta variável e foram superiores ao grupo 4, isto é, os grupos mais precoces foram superiores ao grupo mais tardio para esta variável. Resultado que

corroborar com o de Santiago et al. (2019) que identificou que em áreas férteis o crescimento da soja atingiu melhores produtividades, entretanto acarretou no aumento do acamamento principalmente nas cultivares mais tardias.



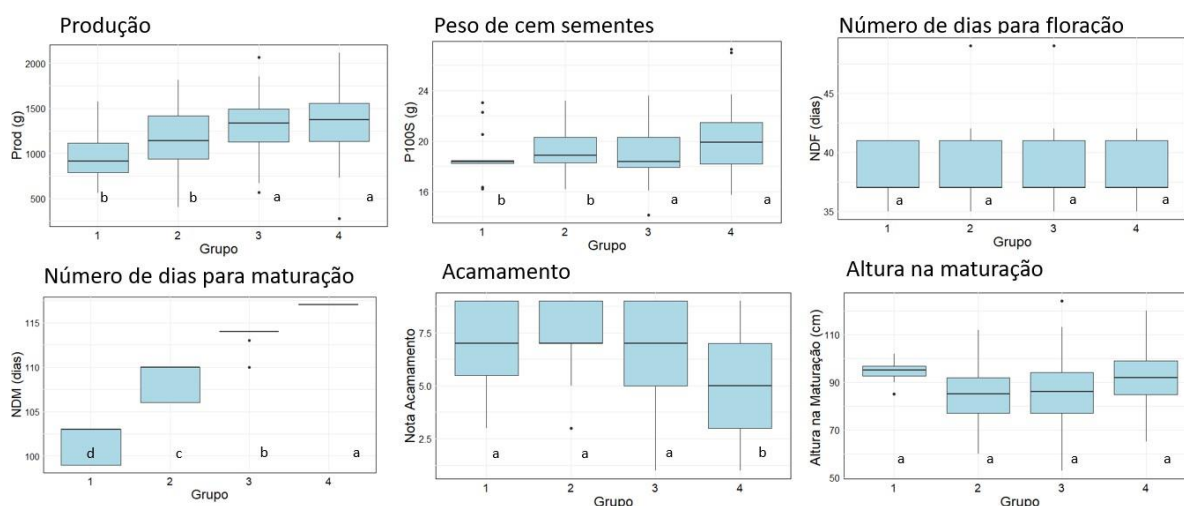
**Figura 10.** Boxplot das características agrônômicas, produção, peso de 100 sementes, acamamento, número de dias para floração (NDF), número de dias para maturação (NDM), e altura de planta na maturação, grupos 1 a 4, população 1.

O grupo 1 da população 2, apresentou maior amplitude entre quartil para as variáveis produtividade e peso de 100 sementes, indicando maior variabilidade dos dados (Figura 11)., para PROD os grupos 3 e 4 foram considerados iguais e apresentaram médias superiores aos grupos 1 e 2 que também tenderam a apresentar médias iguais.

Assim como na população 1, o acamamento foi a variável que apresentou maior amplitude, demonstrando que em todos os grupos foram avaliadas parcelas acamadas (nota 1) e eretas (nota 9), e também assim como na população 1, o grupo 4 apresentou menor média em relação aos demais, ou seja, o grupo mais tardio apresentou maior quantidade de parcelas com menores notas de acamamento, isto é, mais parcelas acamadas.

O grupo 1 apresentou menor amplitude entre quartil em relação aos demais grupos para NDF, indicando que esse grupo apresentou a menor variação para início do florescimento, para esta variável todos os grupos tenderam a ser diferentes em relação a média, onde o grupo 4 foi o que tendeu a apresentar maior média e o grupo 1 a menor.

Para a característica NDM, o grupo 1 foi o mais precoce e o 4 o mais tardio, o qual apresentou a menor amplitude entre os quartis.



**Figura 11.** Boxplot das características agrônômicas, produtividade, peso de 100 sementes, acamamento, número de dias para floração (NDF), número de dias para maturação (NDM) e altura de planta na maturação grupos 1 a 4, população 2.

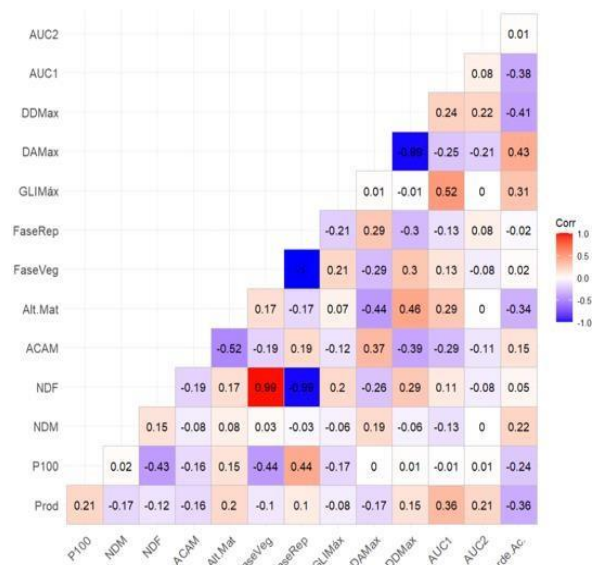
### 3.3 Construção do ideótipo e estudo do comportamento dos genótipos por data de maturação

Existem vários estudos que se mostraram eficientes utilizando fenotipagem de alto rendimento para a construção de um ideótipo de soja. Um exemplo é o trabalho de Roth et al (2020), que propôs um ideótipo que visava a maximização do rendimento mantendo o teor de proteína, por meio do vigor inicial e fechamento de dossel precoce, aumentando o máximo final do índice de área foliar, sem afetar na data de maturação.

No presente estudo o ideótipo foi construído para cada grupo de cada população, com objetivo de obter maior ganho com a seleção para a produtividade, considerando os resultados de correlação de Pearson entre as variáveis da curva de maturação e variáveis agrônômicas. Em caso de  $r \neq 0$ , a correlação foi interpretada segundo Dancy e Reidy (2005),  $r = \pm 0.10 - \pm 0.30$  correlação fraca;  $r = \pm 0.40 - \pm 0.60$  correlação moderada e  $r = \pm 0.70 - \pm 1.00$  correlação forte.

As correlações entre características dentro de cada grupo para as diferentes populações são apresentadas nas Figuras 12 e 13.

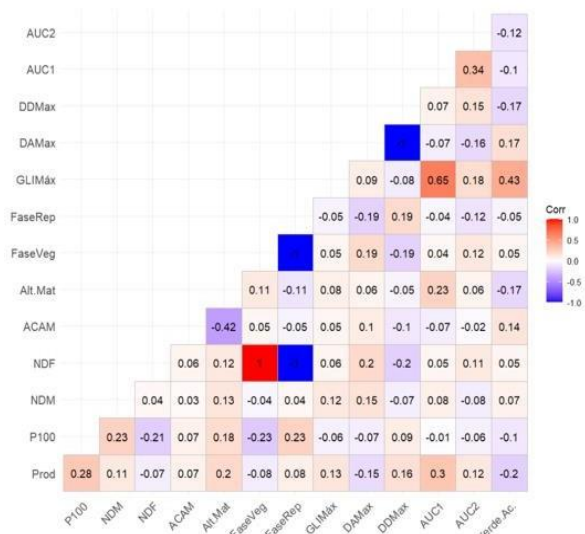
## Grupo 1



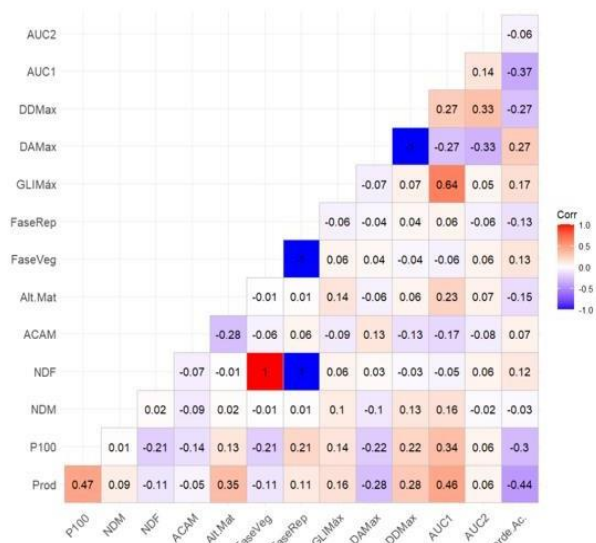
## Grupo 2



## Grupo 3

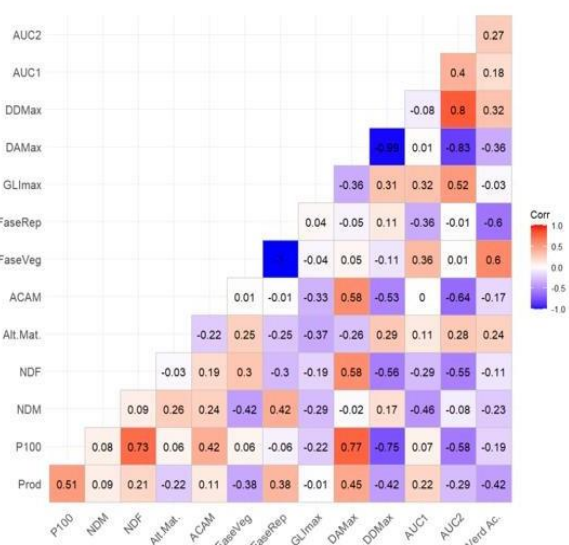


## Grupo 4



**Figura 12.** Heatmap de correlação entre as características agrônômicas e da curva de maturação dos grupos 1 a 4, população 1. Produção (Prod), peso de cem sementes (P100), número de dias para maturação (NDM), número de dias para floração (NDF), acamamento (ACAM), altura na maturação (Alt.Mat), período da fase vegetativa (FaseVeg), período da fase reprodutiva (FaseRep), valor máximo para o índice de vegetação GLI (GLIMax), área total abaixo da curva (AUC 1), área em baixo da curva a partir do valor máximo do índice de vegetação GLI (AUC 2), verde acumulado (Verde.Ac.), número de dias até o valor máximo do índice de vegetação GLI (DAMax) e número de dias depois do valor máximo do índice de vegetação GLI (DDMax).

Grupo 1



Grupo 2



Grupo 3



Grupo 4



**Figura 13.** Heatmap de correlação entre as características agrônômicas e da curva de maturação dos grupos 1 a 4, população 2. Produção (Prod), peso de cem sementes (P100), número de dias para maturação (NDM), número de dias para floração (NDF), acamamento (ACAM), altura na maturação (Alt.Mat), período da fase vegetativa (FaseVeg), período da fase reprodutiva (FaseRep), valor máximo para o índice de vegetação GLI (GLIMáx), área total abaixo da curva (AUC 1), área em baixo da curva a partir do valor máximo do índice de vegetação GLI (AUC 2), verde acumulado (Verde.Ac.), número de dias até o valor máximo do índice de vegetação GLI (DAMax) e número de dias depois do valor máximo do índice de vegetação GLI (DDMax)

Observando o comportamento entre os grupos mais tardios e mais precoces das duas populações em relação as variáveis DAMáx e DDMáx, o único grupo que apresentou correlação positiva com DAMáx foi o grupo 1 da população 2, tal grupo é o mais precoce entre os 8 grupos das duas populações, demonstrando que para grupos precoces é mais favorável que o valor máximo de refletância de verde seja próximo da fase de enchimento de grãos, diferente do observado para os grupos mais tardios.

As variáveis da curva de maturação no presente trabalho foram utilizadas de forma exploratória. Por exemplo, a área embaixo da curva é amplamente utilizada na área de fitopatologia como “área embaixo da curva de progresso de doença (AUDPC)” cujo objetivo é combinar várias observações do progresso da doença em um único valor (SIMKO e PIEPHO, 2012). Simões (2022) utilizou a AUDPC dos valores de refletância dos índices de vegetação e bandas multiespectrais ao longo do tempo das avaliações de ferrugem alaranjada e marrom em cana-de-açúcar, e obteve correlações significativas entre as notas de infecção e os dados de refletância.

Os diferentes resultados das correlações entre produção e as variáveis da curva de maturação, nos diferentes grupos dentro das populações, demonstraram a complexidade das interações entre o desenvolvimento da planta e a produção. Contudo, Carpentieri-Pípolo et al. (2005) reforçaram que a correlação entre variáveis é crucial para uma análise abrangente das relações fenotípicas. Neste sentido é importante relatar que, em programas de melhoramento genético, conhecer as interações entre características é de grande importância para o processo de seleção de genótipos superiores. Assim, procedeu-se o estudo da correlação entre características dentro de cada grupo com o objetivo de definição dos ideótipos para promover a seleção.

### **3.3.1 População 1**

#### **3.3.1.1 Grupo 1**

Neste grupo a produção obteve correlação positiva com P100 ( $r = 0.29$ ), NDM ( $r = 0.44$ ), NDF ( $r = 0.13$ ), ACAM ( $r = 0.24$ ) e com altura na maturação ( $r = 0.21$ ). Esses valores demonstram que, para o grupo 1, os genótipos mais tardios, que demoraram mais para florescer e que não acamaram foram os mais produtivos. Além disso a produção obteve correlação negativa com verde acumulado ( $r = - 0.42$ ), GLIMáx ( $r = - 0.13$ ) e com DAMáx ( $r = - 0.12$ ) indicando que os genótipos mais produtivos possuem tendência de ter menor acúmulo de verde do índice GLL, com

menor número de dias até o valor máximo de refletância de verde e um ponto máximo da curva de maturação mais baixo.

P100 apresentou correlação negativa com NDF ( $r = -0.43$ ) e com altura na maturação ( $r = -0.21$ ), e assim como produtividade apresentou correlação negativa com algumas das variáveis da curva de maturação. P100 também apresentou correlação positiva com a duração da fase reprodutiva ( $r = 0.43$ ).

Neste sentido, observou-se que as cultivares deste grupo que apresentaram uma proporção do ciclo maior para a fase reprodutiva em relação a vegetativa, responderam com um maior peso de cem sementes e tempo de senescência mais longo, tal fato pode ser explicado com maior tempo da decomposição e conversão de clorofila, macromoléculas e proteínas em nutrientes exportáveis que podem ser translocados para as sementes (BHATLA e LAL, 2019).

Para este grupo também foi averiguado que quanto maior o número de dias para floração as plantas tendem a ser mais altas ( $r = 0.42$ ) aumentando a possibilidade de acamamento ( $r = -0.24$ ). Os genótipos que possuíam maior NDF também apresentaram área abaixo da curva de maturação e o GLIMáx maiores, e como NDF obteve correlação negativa com o verde acumulado ( $r = -0.22$ ), conclui-se que, para esse grupo, quando aumenta a fase vegetativa há uma tendência em reduzir o valor do acúmulo de verde.

### **3.3.1.2 Grupo 2**

Produção correlacionou positivamente com AUC 1 ( $r = 0.36$ ), P100 ( $r = 0.21$ ), altura na maturação ( $r = 0.2$ ), com duração da fase reprodutiva ( $r = 0.1$ ), DDMáx ( $r = 0.15$ ) e AUC 2 ( $r = 0.21$ ). Esta variável também apresentou correlações negativa com NDM ( $r = -0.17$ ), NDF ( $r = -0.12$ ), ACAM ( $r = -0.16$ ), fase vegetativa ( $r = -0.1$ ), e verde acumulado ( $r = -0.36$ ). Neste sentido, para este grupo as plantas que produziram mais foram as mais precoces, altas, com menor ciclo, maior proporção da fase reprodutiva em relação à vegetativa, obtiveram um menor acúmulo de verde, e obtiveram uma senescência mais lenta.

A característica P100 apresentou correlação negativa com a fase vegetativa ( $r = -0.44$ ), GLIMáx ( $r = -0.17$ ) e verde acumulado ( $r = -0.24$ ) e correlação positiva com a fase reprodutiva ( $r = 0.44$ ), mostrando que para aumentar o P100 a fase vegetativa deste grupo deveria ser reduzida. Além disso houve uma tendência

de que os genótipos que apresentavam maiores valores de P100 possuíam acúmulo de verde menor.

A correlação negativa de acamamento e altura de plantas ( $r = -0.52$ ), condiz com o estudo de Braccini (2004) sobre época de semeadura de soja, que observou que a semeadura tardia de soja favoreceu o crescimento das plantas, obtendo-se plantas mais altas, porém mais suscetíveis ao acamamento.

### **3.3.1.3 Grupo 3**

Em comparação com os demais grupos deste ambiente, este foi o grupo que apresentou a maior quantidade de variáveis não correlacionadas e também as correlações mais baixas.

Assim como no grupo anterior produção apresentou correlação positiva com P100 ( $r = 0.28$ ), altura na maturação ( $r = 0.2$ ), AUC 1 ( $r = 0.3$ ), AUC 2 ( $r = 0.12$ ) e DDMáx ( $r = 0.16$ ) e negativa com verde acumulado ( $r = -0.2$ ). A correlação entre ACAM e AltMat ( $r = -0.42$ ) também foram condizentes à apresentada para o grupo 2 deste ambiente. Neste sentido, os genótipos mais produtivos tendem a possuir um maior peso de cem sementes, reduzirem o número de dias até o valor de máxima refletância de verde, serem mais tardios, e possuírem uma senescência mais lenta.

Peso de 100 sementes apresentou correlação positiva com NDM ( $r = 0.23$ ), com a duração da fase reprodutiva ( $r = 0.23$ ), e correlação negativa com a duração da fase vegetativa ( $r = -0.23$ ), neste sentido, observamos que para obter maior P100 o ideal seria aumentar o ciclo, estender a fase reprodutiva, ou seja, obter um florescimento mais precoce.

### **3.3.1.4 Grupo 4**

Este grupo foi o que apresentou as maiores correlações entre produção e as variáveis extraídas da curva de maturação sendo elas PROD e AUC 1 ( $r = 0.46$ ), PROD e GLIMáx ( $r = 0.16$ ), PROD e DAMáx ( $r = -0.28$ ), PROD e DDMáx ( $r = 0.28$ ) e PROD e Verde acumulado ( $r = -0.44$ ), apesar da baixa correlação com GLIMáx ( $r = 0.16$ ).

A partir dessas correlações verifica-se que o comportamento dos genótipos mais produtivos desse grupo em relação as variáveis da curva de maturação foi de obterem maior AUC 1, que está correlacionada com NDM ( $r = 0.16$ ), pois quanto maior o ciclo maior a base da área da curva, e quanto maior o GLIMáx maior a altura dessa área, além disso a correlação negativa com verde acumulado mostra que os

genótipos mais produtivos acumularam menos verde. Esse menor acúmulo de verde pode ser justificado por um maior valor de GLI inicial, e não a um menor valor máximo de GLI, pois GLIMáx e verde acumulado possuem correlação positiva ( $r = 0.17$ ).

Além das altas correlações com as variáveis da curva de maturação, a característica produção também apresentou correlação positiva com P100 ( $r = 0.47$ ) e altura na maturação ( $r = 0.35$ ). Neste sentido apesar de quase não existir correlação entre produção e acamamento, esta última apresentou correlação de  $r = -0.28$  com altura de planta, condizendo com os grupos 2 e 3 desde ambiente e demonstrando que apesar de as plantas mais altas serem mais produtivas, elas têm maior tendência a serem suscetíveis ao acamamento.

Peso de cem sementes também apresentou correlação positiva com AUC 1 ( $r = 0.34$ ), GLIMáx ( $r = 0.14$ ), e negativa com verde acumulado ( $r = -0.3$ ), mostrando que P100 e PROD neste grupo apresentam a mesma tendência em relação as variáveis da curva de maturação.

### **3.3.2 População 2**

#### **3.3.2.1 Grupo 1**

Neste grupo a maior parte das correlações significativas foram negativas: PROD e Fase Vegetativa ( $r = -0.38$ ), PROD e DDMáx ( $r = -0.42$ ), PROD e AUC 2 ( $r = -0.29$ ), PROD e Verde acumulado ( $r = -0.42$ ), P100 e DDMáx ( $r = -0.75$ ), P100 e AUC 2 ( $r = -0.58$ ), NDM e AUC 1 ( $r = -0.46$ ), NDF e DDMáx ( $r = -0.56$ ), NDF e AUC 2 ( $r = -0.55$ ), AltMat e ACAM ( $r = -0.22$ ), AltMat e GLIMáx ( $r = -0.37$ ), ACAM e GLIMáx ( $r = -0.33$ ), ACAM e DDMáx ( $r = -0.53$ ), ACAM e AUC 2 ( $r = -0.64$ ), Fase Reprodutiva e AUC 1 ( $r = -0.36$ ), Fase Reprodutiva e Verde Acumulado ( $r = -0.6$ ), DAMáx e AUC 2 ( $r = -0.83$ ).

A partir dessas correlações, pode-se observar que os genótipos mais produtivos foram os que alcançaram o valor de refletância de verde mais tarde, acumularam menos verde e senesceram de forma mais rápida.

O maior peso de cem sementes também estava associado com essas variáveis além de estar correlacionado de forma positiva com NDF ( $r = 0.73$ ), ACAM ( $r = 0.42$ ) e DAMáx ( $r = 0.77$ ), portanto, retardando a floração desse grupo estaria sendo favorável ao aumento do peso das sementes, desta forma as plantas teriam mais tempo para adquirir porte e acumular biomassa, concentrando nutrientes e substâncias essenciais, para translocar para as sementes na época de enchimento

de grãos. A correlação com acamamento condiz com o trabalho de Cruz et al. (2003), que encontrou perda de peso médio de semente de 20% quando as plantas de trigo acamavam no estágio de antese.

Nem produção nem peso de cem sementes obtiveram correlação com NDM, visto a baixa variabilidade dos genótipos para NDM neste grupo (Figura 12).

### **3.3.2.2 Grupo 2**

A característica produção apresentou correlação positiva com P100 ( $r = 0.2$ ), sugerindo que à medida que o peso de cem sementes aumenta, a produção também tende a crescer. Além disso produção também correlacionou de forma positiva com NDF ( $r = 0.15$ ), AltMat ( $r = 0.18$ ) e AUC 2 ( $r = 0.23$ ), indicando que nesse grupo as plantas mais produtivas foram as mais altas, que demoraram para iniciar florescimento e que tiveram a senescência lenta, destacando a importância do tempo de desenvolvimento vegetativo na determinação do rendimento.

Ambas variáveis apresentaram correlação negativa com acamamento, isto quer dizer, que as plantas que obtiveram maiores valores de produção e P100 foram as que mais acamaram, entretanto, esse resultado pode ser explicado pelo fato de a colheita dessas parcelas não ser mecanizada, pois segundo a EMBRAPA (2011) plantas acamadas podem acarretar perda na colheita de até 15%.

A correlação positiva entre AUC 2 e DDMáx ( $r = 0.44$ ), demonstra que nesse grupo os genótipos que atingiram o valor máximo de refletância de verde em um menor número de dias foram os que tiveram a senescência mais lenta.

A AUC 1 apresentou correlação positiva tanto com GLIMáx ( $r = 0.51$ ) quanto com NDM ( $r = 0.1$ ), e devido os valores dessas correlações pode-se dizer que para este grupo quem teve maior determinante no valor da AUC 1 foi o GLIMáx.

### **3.3.2.3 Grupo 3**

Neste grupo, a produção correlacionou de forma positiva com NDF ( $r = 0.24$ ), GLIMáx ( $r = 0.15$ ), DDMáx ( $r = 0.17$ ), e AUC 1 ( $r = 0.29$ ), de forma negativa com DAMáx ( $r = - 0.16$ ), e não apresentou correlação significativa com as demais variáveis. Neste sentido, podemos concluir que os genótipos mais produtivos deste grupo foram os que tenderam a retardar o florescimento, tiveram um GLIMáx mais distante do final do ciclo, e devido à falta de correlação com AUC 2, a velocidade da senescência não interferiu na produção.

A característica P100, apresentou correlação positiva com a duração da fase vegetativa ( $r = 0.19$ ), e negativa com a duração da fase reprodutiva ( $r = - 0.19$ ), além disso também apresentou correlação positiva com AUC 2 ( $r = 0.14$ ), indicando que o ideal seria que quando o objetivo do programa de melhoramento for aumentar o peso de cem sementes deve-se selecionar genótipos que possuam uma fase reprodutiva maior em relação à vegetativa que a senescência dos materiais seja mais lenta.

#### **3.3.2.4 Grupo 4**

Neste ambiente, o grupo 4 foi o que apresentou menor número de correlações significativas entre as características avaliadas e a produção. A característica produção apresentou correlação negativa com DAMáx ( $r = - 0.22$ ) e positiva com DDMáx ( $r = 0.2$ ), indicando que o florescimento precoce foi favorável para o aumento da produtividade. E assim como no grupo 3 deste ambiente, não foi observado correlação com a velocidade de senescência.

O peso de cem sementes obteve correlação negativa com NDM ( $r = - 0.23$ ) e AltMat ( $r = - 0.16$ ), indicando que os genótipos com maiores valores de P100 foram os mais precoces e baixos dentro do grupo.

A correlação negativa de acamamento com AUC 2 ( $r = - 0.21$ ) demonstrou a tendência de que os genótipos que mais acamaram foram os que tiveram a senescência mais lenta, e os que permaneceram eretos apresentaram a senescência mais rápida.

AUC 2 apresentou correlação negativa com DAMáx ( $r = - 0.63$ ), e positiva com DDMáx ( $r = 0.6$ ), demonstrando que, para esse grupo, os genótipos possuem uma senescência mais lenta.

### **3.4 Seleção de genótipos**

Todas as variáveis agronômicas foram utilizadas no índice MGIDI nas duas populações. No entanto, nem todas as variáveis da curva de maturação foram utilizadas, pois na população 1 AUC 2 não foi significativa, na população 2 GLIMáx não foi significativa, e verde acumulado não foi significativa em ambas as populações (Tabela 3).

Todas as variáveis foram ponderadas igualmente (peso 1), e o sentido da seleção variou conforme a correlação da variável com a produção conforme descrito anteriormente no presente trabalho. Optou-se por selecionar os dez melhores

genótipos de cada um dos grupos em cada população, para padronizar-se a quantidade de genótipos selecionados, tendo em vista que cada grupo era composto por quantidade diferente de genótipos.

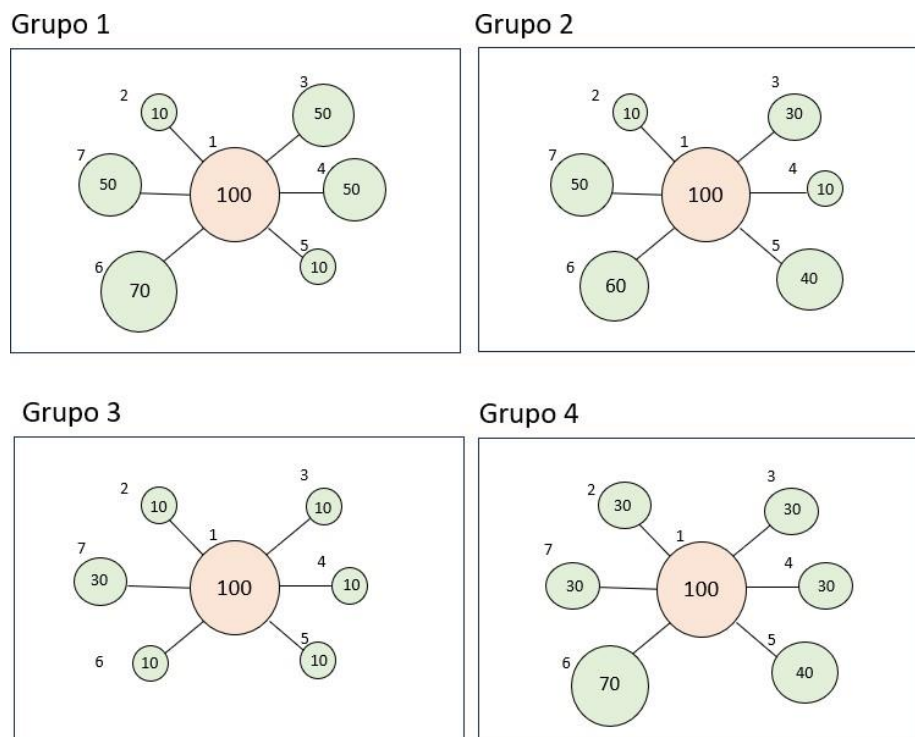
**Tabela 3.** Relação do sentido de seleção das variáveis agronômicas (+ seleção para maiores valores e – seleção para menores valores) e da curva de maturação utilizadas no índice MGIDI, nos grupos 1 a 4 das populações 1 (UEPE Vale da Agronomia) e 2 (UEPE Horta Nova).

População	Grupo	PROD	P100S	Alt.Mat	NDF	NDM	ACAM	GLIMáx	AUC 1	AUC 2	DAMáx	DDMáx
1	1	+	+	+	+	+	+	-	+		-	+
	2	+	+	+	-	-	+	-	+		-	+
	3	+	+	+	+	+	+	+	+		-	+
	4	+	+	+	-	+	+	+	+		-	+
2	1	+	+	-	+	+	+		+	-	+	-
	2	+	+	+	+	+	+		+	+	-	+
	3	+	+	-	+	+	+		+	+	-	+
	4	+	+	+	+	+	+		+	+	-	+

\*PROD: produção, P100S: peso de cem sementes, Alt.Mat: altura de planta na maturação, NDF: número de dias para floração, NDM: número de dias para maturação, ACAM: acamamento, GLIMáx: valor máximo do índice de vegetação GLI, AUC 1: área total abaixo da curva, AUC 2: área em baixo da curva a partir do valor máximo do índice de vegetação GLI, DAMáx: número de dias até o valor máximo do índice de vegetação GLI, DDMáx: número de dias depois do valor máximo do índice de vegetação GLI (DDMax).

### 3.4.1 População 1

Os valores de coincidência entre os 10 melhores genótipos da população 1, das seleções 2 a 7 em relação à seleção 1, variaram de 10% a 70% de coincidência no grupo 1, 10% a 60% no grupo 2, 10% a 30% no grupo 3 e 30% a 70% no grupo 4 (Figura14).



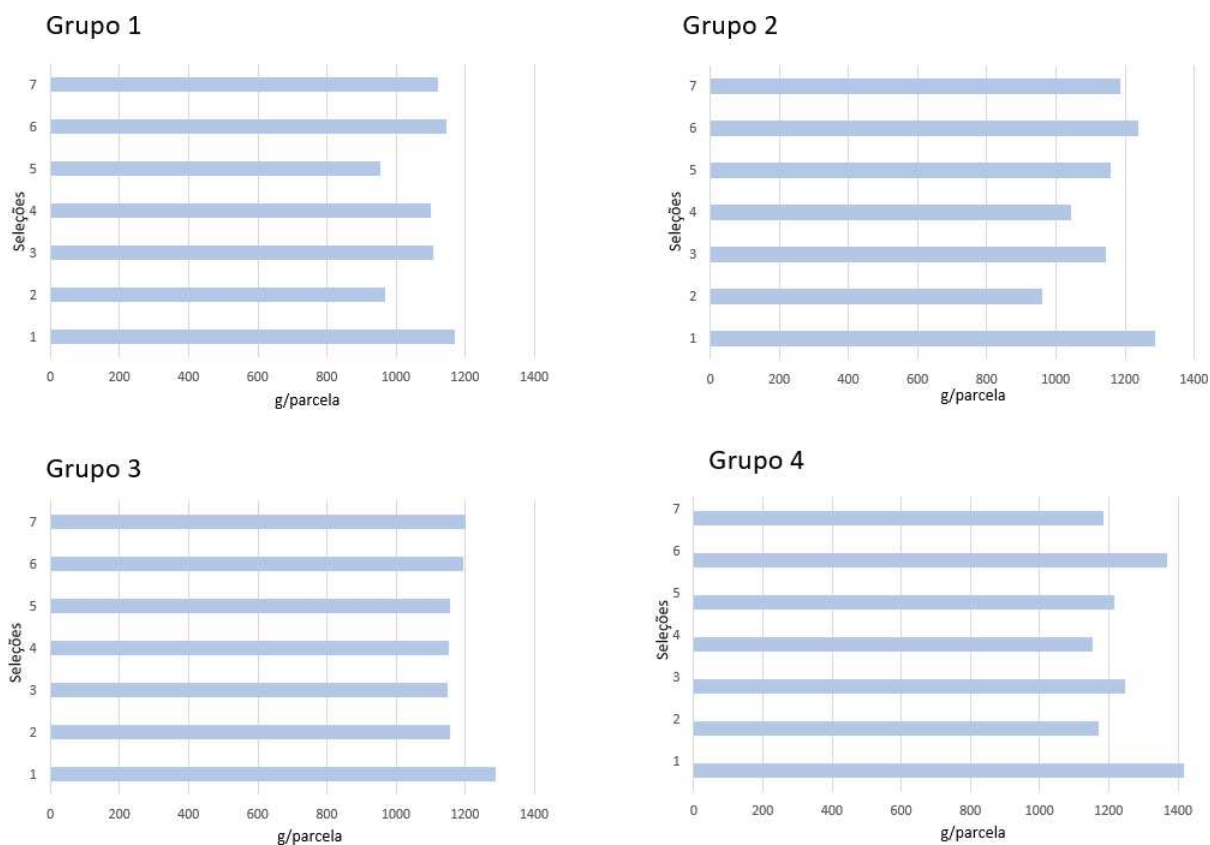
**Figura 14.** Representação gráfica das porcentagens de coincidência entre os dez primeiros genótipos selecionados pelas estratégias 2 a 7 comparados com a estratégia de seleção 1, nos grupos 1 a 4, população 1 (UEPE Vale da Agronomia).

A comunalidade média da população 1 foi superior a 0,5 nas seis seleções nos quatro grupos, sendo todas dentro do valor considerado adequado de no mínimo 0,5 por Matos & Rodrigues (2019).

Para os quatro grupos da população 1, a seleção 2 foi a que apresentou as menores distâncias genótipo - ideótipo variando de 0.29 (genótipo 5273) a 2.46 (genótipo 5022) para o grupo 1, 0.51 (genótipo 5056) a 2.29 (genótipo 5295) grupo 2, 1.36 (genótipo 5062) a 3.10 (genótipo 5027) grupo 3 e 1.08 (genótipo 5057) a 2.86 (genótipo 5286) no grupo 4. A seleção que apresentou as maiores distâncias genótipo-ideótipo foi a seleção 7 que variou de 2.59 (genótipo 5263) a 4.28 (genótipo 5411) grupo 1, 3.60 (genótipo 5091) a 4.42 (genótipo 5524) grupo 2, 2.99 (genótipo 5306) a 4.21 (genótipo 5194) grupo 3 e 2.18 (genótipo 5057) a 3.63 (genótipo 5140) no grupo 4.

A média de peso de parcela total da população foi 1018,37 g/parcela, a média das testemunhas foi 1008,97 g/parcela, média total do grupo 1 foi de 977,79 g/parcela, grupo 2 de 999,70 g/parcela, grupo 3 de 1049,4 g/parcela e do grupo 4 de 1032,87

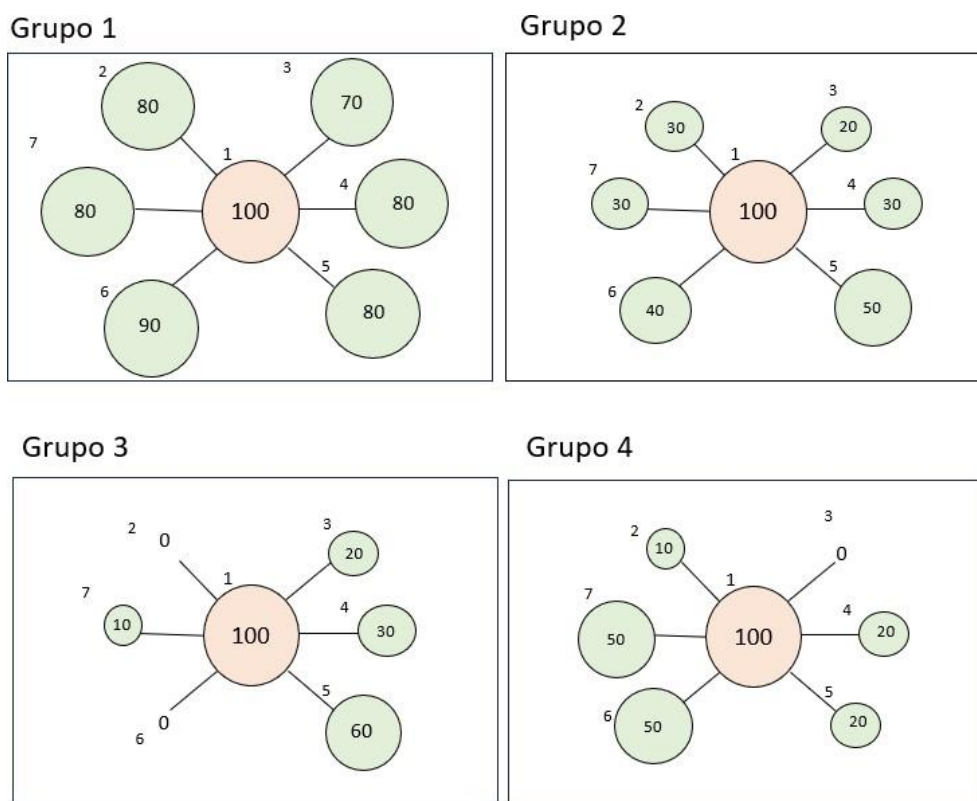
g/parcela. As médias dos dez primeiros selecionados em cada uma das seleções estão apresentadas na Figura 15.



**Figura 15.** Gráfico das médias dos dez melhores genótipos nas Seleções 1 a 7 na População 1 (UEPE Vale da Agronomia).

### 3.4.2 População 2

Na população 2, os valores de coincidência entre os 10 melhores genótipos das seleções 2 a 7 em relação a seleção 1, variaram de 80% a 90% de coincidência no grupo 1, 20% a 50% no grupo 2, 0% a 60% no grupo 3 e 0% a 50% no grupo 4 (Figura 16).



**Figura 16.** Representação gráfica das porcentagens de coincidência entre os dez primeiros genótipos selecionados pelas estratégias 2 a 7 comparados com a estratégia de seleção 1, nos grupos 1 a 4, População 2 (UEPE Horta Nova).

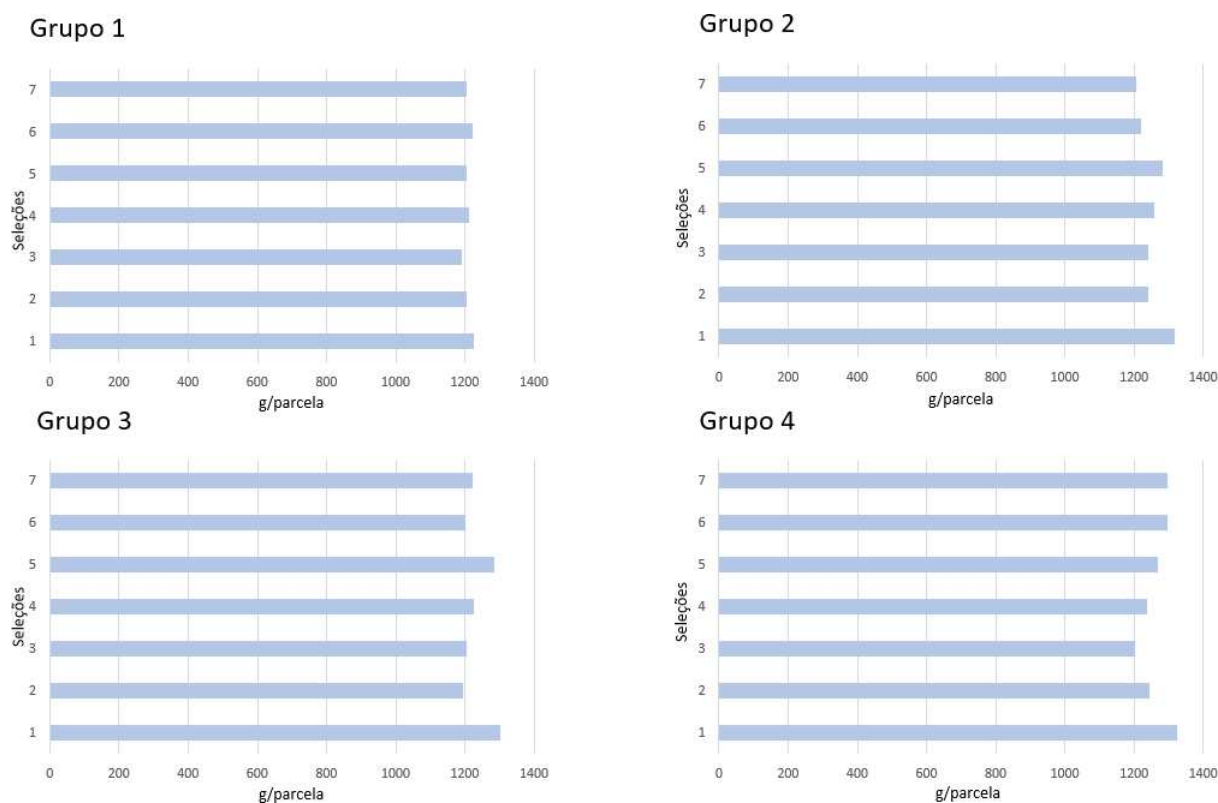
Diferentemente da população 1, a comunalidade média da população 2 para a seleção 6 do grupo 2 e a seleção 2 do grupo 3 foram 0,37 e 0,45 respectivamente, sendo as únicas abaixo do valor considerado adequado de no mínimo 0,5 por Matos & Rodrigues (2019).

Este resultado sugere que, nessas condições específicas, as variáveis associadas às seleções 2 e 6 dos grupos 2 e 3 na população 2 podem não compartilhar uma proporção satisfatória de variância comum, o que pode impactar a eficácia da análise fatorial nesse contexto.

Nesta população, para os grupos 2, 3 e 4 a seleção 2 foi a que apresentou as menores distâncias genótipo – ideótipo variando de 0.70 (genótipo 11276) a 1.79 (genótipo 11186) para o grupo 2, 0.01 (genótipo 11104) a 0.24 (genótipo 11139) grupo 3 e 0.95 (genótipo 11089) a 1.66 (genótipo 11050) grupo 4, já para o grupo 1 as menores distâncias genótipo – ideótipo foram encontradas na seleção 6 e variaram de 0.11 (genótipo 11230) a 2.06 (genótipo 11288).

As maiores distâncias genótipo – ideótipo foram obtidas na seleção 7, para os quatro grupos. No grupo 1, variou de 3.09 (genótipo 11289) a 5.18 (genótipo 11268), grupo 2 variou de 2.71 (genótipo 11201) a 4.06 (genótipo 11180), grupo 3 de 2,41 (genótipo 11211) a 3.19 (genótipo 11240) e no grupo 4 essa distância variou de 2.55 (genótipo 11111) a 3.95 (genótipo 11291).

A média de peso de parcela total da população foi de 1212,50 g/parcela. A média das testemunhas foi de 1212,49 g/parcela e as médias totais dos grupos 1, 2, 3 e 4 foram, respectivamente, 1200,62, 1217,71, 1212,29 e 1208,18 g/parcela. As médias dos dez primeiros selecionados em cada uma das seleções estão na Figura 17.



**Figura 17.** Gráfico das médias dos dez melhores genótipos nas Seleções 1 a 7 na População 2 (UEPE Horta Nova).

### 3.4.3 Ganho com seleção

Assim como esperado, o ganho genético com a seleção direta (Seleção 1) para peso de parcela foi superior às demais seleções nos dois ambientes, tanto em relação à média da população (Tabela 5) quanto em relação à média dos grupos (Tabela 6).

**Tabela 5.** Tabela de porcentagem de ganho de seleção em relação à média da População e em relação à média do Grupo, nas populações 1 (UEPE Vale da Agronomia) e 2 (UEPE Horta Nova).

Ganho em relação à média da População					
População	Seleção	Grupo 1	Grupo 2	Grupo 3	Grupo 4
1	1	10,21	18,13	18,23	26,80
	2	-3,33	-3,81	9,28	10,25
	3	5,88	8,53	8,90	15,48
	4	5,51	1,77	9,13	9,02
	5	-4,41	9,31	9,18	13,19
	6	8,56	14,76	11,90	23,59
	7	6,94	11,30	12,35	11,06
2	1	0,19	1,58	1,35	1,66
	2	-0,10	0,41	-0,26	0,47
	3	-0,31	0,41	-0,11	-0,16
	4	-0,02	0,70	0,19	0,39
	5	-0,10	1,04	1,10	0,85
	6	0,13	0,13	-0,19	1,27
	7	-0,13	-0,13	0,16	1,25

**Tabela 6.** Tabela de porcentagem de ganho de seleção em relação à média do Grupo, nas Populações 1 (UEPE Vale da Agronomia) e 2 (UEPE Horta Nova).

Ganho em relação à média do Grupo					
População	Seleção	Grupo 1	Grupo 2	Grupo 3	Grupo 4
1	1	13,48	19,75	15,67	25,46
	2	-0,62	-2,60	6,98	9,15
	3	8,97	9,97	6,61	14,30
	4	8,58	3,09	6,83	7,93
	5	-1,74	10,77	6,88	12,04
	6	11,77	16,32	9,52	22,30
	7	10,08	12,80	9,96	9,95
2	1	0,37	1,49	1,35	1,73
	2	0,08	0,33	-0,26	0,53
	3	-0,13	0,33	-0,11	-0,10
	4	0,16	0,62	0,20	0,46
	5	0,08	0,95	1,10	0,92
	6	0,31	0,05	-0,18	1,34
	7	0,05	-0,20	0,16	1,32

Verifica-se que na população 2 os ganhos genéticos foram inferiores em relação aos apresentados na população 1. A diferença no ganho genético com a seleção entre os dois ambientes pode ser explicada pela discrepância observada da herdabilidade de peso de parcela. Na população 1 a herdabilidade foi de 68,67%, enquanto que na população 2 esta foi de 18%. Conforme mencionado por Cruz et. al (2012), menores valores de herdabilidade estão associados a menores ganhos com a

seleção, e os valores de herdabilidade estão correlacionados com a diferença de populações que originou os genótipos de cada ambiente. Os materiais da população 1 foram originados de 84 populações e da população 2 foram originados de 15 populações.

A seleção 2, é uma solução eficaz para situações onde existe uma grande quantidade de parcelas a serem avaliadas e selecionadas, mas a mão de obra é limitada ou não possui capacitação para vencer a demanda de trabalho de forma eficiente e rápida. Neste sentido, utilizar imagens de drone para coletar informações fenotípicas auxilia a selecionar genótipos superiores com uma maior precisão.

As seleções 5 e 7, não são viáveis aos programas de melhoramento genético se o objetivo for apenas ganho em produção. Ambas são inferiores à seleção direta para peso de parcela. O método de seleção 7 obteve menores ganhos em relação a seleção 6.

A seleção 6 se mostrou-se eficiente nos quatro grupos da população 1, e nos grupos 1, 2 e 4 da população 2, apresentando ganho equilibrado para todas as variáveis de interesse. Sendo recomendado seu uso quando o interesse do programa for selecionar mais de uma variável de forma simultânea.

#### **3.4.3.1 População 1**

A seleção indireta através das variáveis da curva de maturação (Seleção 2), se mostrou satisfatória em relação à média total do grupo e em relação à média da população, nos grupos 3 (+7 % e +9,3%) e 4 (+9 % e + 10,2%). Entretanto, nos grupos 1 e 2 esta seleção não foi positiva, em relação à média dos grupos (-0,62% e -2,60 %, respectivamente) nem em relação à média da população (-3,3% e -3,8%, respectivamente), tal fato pode ser explicado devido a incidência de ferrugem neste ensaio no início da maturação, prejudicando principalmente os genótipos mais precoces.

A seleção indireta através das variáveis agronômicas (Seleção 3), se mostrou satisfatórias nos quatro grupos desta população, apresentando ganhos de 5,8 % grupo 1, 8,5% grupo 2, 8,9% grupo 3 e 15,4% grupo 4 em relação à média da população, e ganhos de 8,9 % grupo 1, 9,9% grupo 2, 6,6% grupo 3 e 14,3% grupo 4 em relação à média do grupo. Destacando-se como a seleção indireta mais eficiente para esta população.

A seleção 4, não seria viável na população 1, pois nos grupos 1, 2, e 4 esta seleção foi inferior à seleção 3, e no grupo 3 foi inferior à seleção 2. Neste sentido

seria preferível selecionar os genótipos pelas variáveis agronômicas ou pelas variáveis da curva de maturação, conforme o objetivo do programa de melhoramento.

Nesta população, o ranqueamento das estratégias de seleções em relação à média da população e em relação à média do grupo variou entre os grupos (Tabela 7). No entanto, em geral na população 1, o ranqueamento das estratégias foi: seleção 1, seleção 6, seleção 7, seleção 5, seleção 3, seleção 4 e seleção 2.

Apesar de o ranqueamento variar entre os grupos, para os grupos 1, 2 e 4 o ranqueamento das estratégias foi o mesmo tanto em relação à média da população quanto em relação à média do grupo. Entretanto, no grupo 3 houve mudança no ranqueamento das estratégias de seleção, demonstrando a importância de separar os genótipos por grupos conforme o ciclo.

A partir deste ranqueamento conclui-se que a utilização das características agronômicas foi eficiente e que a inclusão de características da curva de maturação na seleção 7 foram importantes na promoção de ganho com a seleção.

**Tabela 7.** Tabela de ranqueamento das estratégias de seleção conforme o ganho de seleção em relação à média da população e em relação à média do grupo, na População 1 (UEPE Vale da Agronomia).

Ranqueamento das estratégias- ganho em relação à média da população				
População	Grupo 1	Grupo 2	Grupo 3	Grupo 4
1	1	1	1	1
	6	6	7	6
	7	7	6	3
	3	5	5	5
	4	3	4	7
	2	4	2	2
	5	2	3	4
Ranqueamento das estratégias- ganho em relação à média dos grupos				
População	Grupo 1	Grupo 2	Grupo 3	Grupo 4
1	1	1	1	1
	6	6	7	6
	7	7	6	3
	3	5	2	5
	4	3	5	7
	2	4	4	2
	5	2	3	4

### 3.4.3.2 População 2

Nesta população, para a seleção indireta através das variáveis da curva de maturação (Seleção 2), o grupo 1 se comportou de forma divergente quanto ao ganho

em relação à média do grupo (+ 0,08%) e média da população (- 0,09%). Os grupos 2 e 4 apresentaram ganhos tanto em relação à média do grupo +0,3% e +0,5% respectivamente, quanto em relação à média da população, +0,4 % em ambos os grupos. Nesta população, o grupo 3 foi o único que obteve ganho negativo em relação ao grupo (-0,25%) e em relação a população (-0,26%).

A seleção indireta através das variáveis agronômicas (Seleção 3), não foi eficiente na população 2. Apresentando ganho apenas no grupo 2, sendo estes de 0,41% em relação à média da população e 0,33% em relação à média do grupo. Nos grupos 1 e 4, esta foi a única seleção que obteve valor negativo para ganho em relação à média do grupo ( -0,13% e -0,09%), apesar do grupo 3 também apresentar redução na média desta variável em relação ao grupo (- 0,10 %), as seleções 2 e 6 (- 0,18%) também não foram satisfatórias neste grupo. Em relação à média populacional o ganho obtido foi de - 0,31% grupo 1, - 0,11% grupo 3 e - 0,16 % no grupo 4.

A seleção que utilizou as variáveis agronômicas e as variáveis da curva de maturação (Seleção 4), foi superior as seleções 2 e 3, com exceção do grupo 4, onde a seleção 2 foi superior a seleção 4, podendo ser utilizada em situações onde a seleção deve ser feita antes da colheita dos genótipos.

Nesta população, o ranqueamento das estratégias de seleções em relação à média da população e em relação à média do grupo variou entre os grupos (Tabela 8). Mas em geral na população 2, o ranqueamento das estratégias foi: seleção 1, seleção 5, seleção 6, seleção 4, seleção 7, seleção 2 e seleção 3.

Apesar de o ranqueamento variar entre os grupos, o ranqueamento das estratégias para os quatro grupos foi o mesmo tanto em relação à média da população quanto em relação à média do grupo. Assim como na população 1, demonstrou-se a importância de se selecionar dentro de grupos conforme o ciclo.

**Tabela 8.** Tabela de ranqueamento das estratégias de seleção conforme o ganho de seleção em relação à média da população e em relação a média do grupo, na População 2 (UEPE Horta Nova).

Ranqueamento das estratégias- ganho em relação à média da população				
População	Grupo 1	Grupo 2	Grupo 3	Grupo 4
	1	1	1	1
	6	5	5	6
	4	4	4	7
2	5 e 2	2 e 3	7	5
	7	6	3	2
	3	7	6	4
			2	3
Ranqueamento das estratégias- ganho em relação à média dos grupos				
	Grupo 1	Grupo 2	Grupo 3	Grupo 4
	1	1	1	1
	6	5	5	6
	4	4	4	7
2	5 e 2	2 e 3	7	5
	7	6	3	2
	3	7	6	4
			2	3

#### 4. CONCLUSÃO

A associação das características agronômicas e características da curva de maturação foram eficientes para definir o ideótipo de cada grupo em cada população e selecionar genótipos superiores em cada um dos grupos a partir do ideótipo definido, nas diferentes estratégias de seleção.

## 5. REFERÊNCIAS

ALMEIDA, L.A.; KIIH, R.A.S.; MIRANDA, M.A.C.; CAMPELO, G.J.A. **Melhoramento da soja para regiões de baixas latitudes. Recursos Genéticos e Melhoramento de Plantas para o Nordeste Brasileiro.** Março, 1999.

APPELHANS, T.; DETSCH, F.; REUDENBACH, C.; WOELLAUER, S. **mapview: Interactive viewing of spatial data in R.** In: EGU General Assembly 2016, held 17-22 April, 2016 in Vienna Austria, id. EPSC2016-1832. Publicado em abril de 2016.

ARAUS, J.L., and J.E. CAIRNS. **Field high-throughput phenotyping: the new crop breeding frontier.** *Trends Plant Sci.* 19(1): 52–61. 2018. DOI: 10.1016/j.tplants.2013.09.008.

ARAUS, J. L., KEFAUVER, S. C., ZAMAN-ALLAH, M., OLSEN, M. S., & CAIMS, J. E. **Translating High-Throughput Phenotyping into Genetic Gain.** *Trends in Plant Science*, 23(5), 451-466.2018. DOI: 10.1016/j.tplants.2018.02.001.

BATES, D.; M, M.; BOLKER, B. M.; WALKER, S. C. **Fitting linear mixed-effectes model using lme4.** *Journal of Statistical Software.* v. 67, 2015.

BHATLA, S.C.; LAL, M. A. **Plant Physiology, Development and Metabolism.** 1. ed. Singapore: Springer Singapore, 2019. XXXIV, 1237 p. DOI: 10.1007/978-981-13-2023-1.

BERNARDI, A. d. C. et al. **Agricultura de precisão: resultados de um novo olhar.** Brasília, DF: Embrapa, 596 p., 2014.

BRACCINI, A. DE L. E et al. **Características agronômicas e rendimento de sementes de soja na semeadura realizada no período de safrinha.** *Bragantia*, v. 63, n. 1, p. 81–92, 2004.

BORÉM, A.; MIRANDA, G. V.; FRITSCHÉ-NETO, R. **Melhoramento de plantas.** 7. ed. Viçosa, MG: Editora UFV, 2017.

CASAGRANDE, C.R. **Associação entre fenotipagem de alto rendimento do dossel e produtividade em soja**. 2019. 50p. Dissertação em Genética e Melhoramento -Universidade Federal de Viçosa, Viçosa/MG.

CARPENTIERI-PÍPOLO, V.; GASTALDI, L. F.; PIPOLO, A. E. **Correlações fenotípicas entre caracteres quantitativos em soja**. *Semina: Ciências Agrárias, Londrina*, v. 26, n. 1, p. 11-16, 2005.

CLEVELAND, W. S. **Robust locally weighted regression and smoothing scatterplots**. *J. Amer. Statist. Assn.* 74, p. 829-836, 1979

CONAB. **Boletim da safra de Grãos**. Disponível em: <https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos>. Acesso em 29/10/2023.

CRUZ, C.D.; REGAZZI, A.J.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. 4.ed. Viçosa: Ed. da UFV, 2012. p. 211-221.

CRUZ, P. J.; CARVALHO, F. I. F. de; SILVA, S. A.; KUREK, A. J.; BARBIERI, R. L.; CARGNIN, A. **Influência do acamamento sobre o rendimento de grãos e outros caracteres em trigo**. *Revista Brasileira de Agrociência*, v. 9, n. 1, p. 05-08, jan-mar, 2003.

DANCEY, Christine & REIDY, John. **Estatística Sem Matemática para Psicologia: Usando SPSS para Windows**. PortoAlegre, Artmed, 2006.

DE FELIPE, M., GERDE, J. A., E ROTUNDO, J. L. **Soybean genetic gain in maturity groups III to V in Argentina from 1980 to 2015**. *Crop science*, v. 56, n. 6, p. 3066-3077, 2016

DONALD, C. M. **In search of yield**. *J. Aust. Inst. agric. Sci.* 28: 171–178. 1962.

DUNN, O. J. **Multiple comparisons using rank sums**. *Technometrics*, 6(3):241-252. 1964.

EMBRAPA. **Comunicado Técnico 271:Perdas na colheita na cultura da soja.** Pelotas, RS: dezembro, 2011. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/79567/1/comunicado-271.pdf>. Acesso em: 09/10/2023.

EMBRAPA. **Empresa Brasileira de Pesquisa Agropecuária. Centro Nacional de Pesquisa de Soja. Tecnologia de produção de soja -região central do Brasil 2015.** Londrina: Embrapa Soja -Embrapa Cerrados -Embrapa Agropecuária Oeste, 2014, 265 p.

FIORANI, F., and SCHURR, U. **Future scenarios for plant phenotyping.** Annu. Rev. Plant Biol. 64:267–291, 2013.

GREGERSEN, P.L., CULETIC, A., BOSCHIAN, L. et al. **Plant senescence and crop productivity.** Plant Mol Biol **82**, 603–622 (2013). <https://doi.org/10.1007/s11103-013-0013-8>.

HIJIMANS R.J. **terra: spatial data analysis.** (2020). Disponível em: <https://CRAN.R-project.org/package=terra>.

KRUSKAL, W.H.; WALLIS, W.A. **Use of ranks in on-criterion variance analyses.** Journal of the American Statistical Association, v.47, n.260, p.583-621, 1952.

LI, J., and Y. SHI. 2018. **Elucidating Sorghum Biomass, Nitrogen and Chlorophyll Contents With Spectral and Morphological Traits Derived From Unmanned Aircraft System.** 9 (October): 1–12. doi:10.3389/fpls.2018.01406.

LOUHAICHI, M.; BORMAN, M. M.; JOHNSON, D. E. **Spatially located platform and aerial photography for documentation of grazing impacts on wheat.** Geocarto International, v. 16, n. 1, p. 65–70, mar. 2001.

MASUDA, T.; GOLDSMITH, P. D. **World soybean production: area harvested, yield, and long-term projections.** International food and agribusiness management review, v. 12, n. 1030-2016-82753, p. 1-20, 2009.

MEDINA, G., G. RIBEIRO & E.M. Brasil. 2016. **Participação brasileira na cadeia da soja:881 lições para o futuro do agronegócio nacional**. R. de Econ. e Agronegócio, 13: 4-9.

MUNIZ, F. R. S. **Análise da variabilidade genética em populações segregantes de soja**. 2007. 94p. Tese de Doutorado em Agronomia-Universidade Estadual Paulista, Jaboticabal/SP.

MATIAS F.I, CARAZZA-HARTER M.V, ENDELMAN J.B. **FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials**. The Plant Phenome J. 2020. DOI: <https://doi.org/10.1002/ppj2.20005>.

MATOS, D. A. S.; RODRIGUES, E. C. **Análise Fatorial**. 1 ed. Brasília: Enap, 2019. 74 p.

OLIVOTO, T., and LÚCIO, A.D. **Metan: an R package for multi-environment trial analysis. Methods**. Ecol Evol. 11:783-789, 2020. doi:10.1111/2041-210X.13384

OLIVOTO, T.; NARDINO, M. **MGIDI: Toward an effective multivariate selection in biological experiments**. Bioinformatics, v. 37, n. 10, p. 1383-1389, 2021.

PAWAR P & MATIAS F.I. **FIELDimageR.Extra: Advancing user experience and computational efficiency for analysis of orthomosaic from agricultural field trials**. The Plant Phenome J. 2023;

PEBESMA, E. **Simple Features for R: Standardized Support for Spatial Vector Data**. The R Journal, v. 10, n. 1, p. 439-446, 2018.

PEBESMA, E. et al. **stars: Spatiotemporal Arrays, Raster and Vector Data Cubes**. Versão 0.6-4. 2023. Disponível em: <https://cran.r-project.org/package=stars>.

R Core Team. **stats: The R Stats Package**. Versão 4.3.1. R Foundation for Statistical Computing, Viena, Áustria, 2023. Disponível em: <https://cran.r-project.org/package=stats>.

RAY, D. K. MUELLER, N. D., WEST, P. C., & FOLEY, J. A. **Yield trends are insufficient to double global crop production by 2050.** PloS one, v. 8, n. 6, p. e66428, 2013.

ROTH, L., BARENDREGT, C., BÉTRIX, C.-A., HUND, A., & WALTER, A. **High-throughput field phenotyping of soybean: Spotting an ideotype.** Remote Sensing of Environment, 269, 112797. 2022.

SAGAE, V.S. **Mapeamento de herdabilidade regional para resposta de cultivares de soja (glycine max l.) submetidas ao deficit hídrico.** 2022. 54 f. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa/MG. 2022

SANTIAGO, A. D., PROCÓPIO, S. de O., CARVALHO, H. W. de L., BRAZ, G. B. P. **Desempenho de Cultivares de Soja em Áreas com Histórico de Produção de Cana-de-Açúcar no Sealba.** Boletim de Pesquisa e Desenvolvimento, Embrapa Tabuleiros Costeiros, Aracaju, SE, Brasil. Novembro de 2019. ISSN 1678-1961.

SARKAR, S.; ZHOU, J.; SCABOO, A.; ZHOU, J.; ALOYSIUS, N.; LIM, T.T. **Assessment of Soybean Lodging Using UAV Imagery and Machine Learning.** *Plants* 2023, 12, 2893. DOI: <https://doi.org/10.3390/plants12162893>.

SEDIYAMA, T.; SILVA, F.; BORÉM, A. **Soja do plantio à colheita.** Viçosa, MG: Ed UFV. 2015. 333p.

SMITH, HF. **A discriminant function for plant selection.** Ann. Eugen., v.7, p. 240-250, 1936

SIMKO, I., and PIEPHO, H.-P. **The area under the disease progress stairs: Calculation, advantage, and application.** Phytopathology 102:381-389, 2012.

SIMÕES, I. O. P. S., **Fenotipagem de Alto Rendimento Baseada em Campo para Resistência da Cana-de-Açúcar às Ferrugens Alaranjada e Marrom.** Tese de

Doutorado, Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2022.

SILVA, F. L. DA et al. **Melhoramento da Soja**. Viçosa, MG: Ed. UFV, 2017.

SILVA, F. L. DA et al. **Soja: do plantio à colheita**. 2. ed. São Paulo, SP: Oficina de Textos, 2022. 312p.

USDA- United States Department of Agriculture. Foreign Agricultural Service online disponível em: <[https:// www.fas.usda.gov](https://www.fas.usda.gov)>. Acesso em: 30/01/2024.

VOLPATO, L. **Seleção de progênies de soja para produção de grãos com uso de modelos mistos**. 2016. Dissertação de Mestrado em Genética e Melhoramento - Universidade Federal de Viçosa, Viçosa/MG

VOLPATO, L.; DOBBELS, A.; BOREM, A.; LORENZ, A. J. **Optimization of temporal UAS-based imagery analysis to estimate plant maturity date for soybean breeding**. *Plant Phenome Journal*, [S.l.], v. 4, n. 1, p. e20018, May 2021. DOI: <https://doi.org/10.1002/ppj2.20018>.

YANG, W., GUO, Z., HUANG, C., DUAN, L., CHEN, G., JIANG, N., FANG, W., FENG, H., XIE, W., LIAN, X., et al. (2014). **Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice**. *Nat. Commun.* 5:5087, 2014.