

FERNANDA CUPERTINO RODRIGUES

**ESTRATÉGIAS DE SELEÇÃO DE LINHAGENS DE SOJA QUANTO A
MODELAGEM ESTATÍSTICA E DEFINIÇÃO DE IDEÓTIPO**

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

Orientador: Felipe Lopes da Silva

Coorientadores: Luiz Alexandre Peternelli
Maicon Nardino

**VIÇOSA – MINAS GERAIS
2024**

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

T

R696e
2024

Rodrigues, Fernanda Cupertino, 1994-
Estratégias de seleção de linhagens de soja quanto a
modelagem estatística e definição de ideótipo / Fernanda
Cupertino Rodrigues. – Viçosa, MG, 2024.
1 tese eletrônica (84 f.): il. (algumas color.).

Orientador: Felipe Lopes da Silva.
Tese (doutorado) - Universidade Federal de Viçosa,
Departamento de Agronomia, 2024.
Inclui bibliografia.
DOI: <https://doi.org/10.47328/ufvbbt.2024.538>
Modo de acesso: World Wide Web.

1. Soja - Melhoramento genético. 2. Soja - Seleção.
3. Análise espacial (Estatística). 4. Modelos multiníveis
(Estatística). 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

FERNANDA CUPERTINO RODRIGUES

**ESTRATÉGIAS DE SELEÇÃO DE LINHAGENS DE SOJA QUANTO A
MODELAGEM ESTATÍSTICA E DEFINIÇÃO DE IDEÓTIPO**


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

APROVADA: 27 de junho de 2024.

Assentimento:

Documento assinado digitalmente
 **FERNANDA CUPERTINO RODRIGUES**
Data: 17/08/2024 19:54:00-0300
Verifique em <https://validar.itl.gov.br>

Fernanda Cupertino Rodrigues
Autora

Documento assinado digitalmente
 **FELIPE LOPES DA SILVA**
Data: 25/08/2024 19:43:21-0300
Verifique em <https://validar.itl.gov.br>

Felipe Lopes da Silva
Orientador

A todos que de alguma forma contribuíram para o meu aprendizado.
DEDICO

AGRADECIMENTOS

A Deus, pela graça da vida e por me guiar pelos caminhos que me trouxeram até aqui.

À minha família, pelo apoio compreensão e amor.

Aos meus queridos amigos, que sempre caminham comigo, seja por qual caminho for.

Ao professor Felipe Lopes da Silva.

Ao professor Luiz Alexandre Peternelli, pelo suporte e incentivo.

Ao professor Maicon, pela parceria.

Aos estagiários, funcionários e demais amigos de pós-graduação pelas ideias compartilhadas, discussões técnicas científicas, apoio e amizade.

Ao Programa Soja, pela oportunidade de crescimento pessoal e profissional.

À Universidade Federal de Viçosa, bem como ao Programa de Pós-Graduação em Genética e Melhoramento e ao Departamento de Agronomia pelo acolhimento e pelas oportunidades a mim concedidas e por todo aprendizado ao longo desses anos.

Às agências de fomento: Conselho Nacional de Desenvolvimento Científico e Tecnológico, pela concessão da bolsa de estudos; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, código de financiamento 001; e, Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG.

Nothing is as ours as our dreams.
(Friedrich Nietzsche)

RESUMO

RODRIGUES, Fernanda Cupertino, D.Sc., Universidade Federal de Viçosa, junho de 2024. **Estratégias de seleção de linhagens de soja quanto a modelagem estatística e definição de ideótipo.** Orientador: Felipe Lopes da Silva. Coorientadores: Luiz Alexandre Peternelli e Maicon Nardino

Em um programa de melhoramento genético, a escolha de modelos estatísticos que forneçam valores genotípicos acurados é de extrema importância. Não obstante, índices que fazem uso das várias características de interesse agrônomo para a seleção de genótipos superiores também são uma ferramenta importante para a obtenção de incremento em tais características. Este trabalho foi dividido em dois capítulos e objetivou: (1) Comparar modelos estatísticos, associados às estratégias de análises citadas, que favoreçam a maior eficiência na predição de valores genéticos de linhagens para características agrônomicas de interesse; e (2) Comparar diferentes cenários para construção do ideótipo com o uso do MGIDI para a seleção de linhagens superiores de soja. Para tanto, ensaios foram implantados em Capinópolis nos anos agrícolas de 2020/21 e 2021/22 e em Madre de Deus nos anos agrícolas de 2019/20 e 2021/22. As características avaliadas para os ensaios conduzidos em Capinópolis foram: produção (g), altura de planta (cm) e ciclo (dias), já em Madre de Deus a característica avaliada foi produção (g). No primeiro capítulo três modelos foram testados: (1) Delineamento em Blocos Aumentados (DBA); (2) Análise Espacial autorregressiva de duas dimensões (AR1 x AR1); e (3) Análise Espacial autorregressiva de duas dimensões com adição de bloco. Os componentes de variância foram obtidos pela metodologia de modelos mistos *Restricted Maximum Likelihood* (REML) e os valores genotípicos foram preditos utilizando *Best Linear Unbiased Prediction* (BLUP). Analisou-se cada ambiente separadamente para cada variável. Para comparar os três modelos foram utilizadas a herdabilidade proposta por Cullis (h^2_g), a Acurácia com base na herdabilidade de Cullis (r_{gg}), *Akaike Information Criterion* (AIC) e *Bayesian Information Criteria* (BIC). Concluiu-se que os modelos que melhor se ajustam são os que consideram a correlação espacial entre as parcelas. Para três ambientes é preferencial a utilização da análise espacial sem a adição de blocos e para um ambiente é preferível o modelo que abrange a análise espacial com a adição de blocos. De posse dos valores genotípicos obtidos por meio dos modelos mistos, o MGIDI foi aplicado em diferentes cenários para verificação da melhor estratégia de definição do ideótipo. O ideótipo foi desenhado em seis diferentes

cenários: (1) considerando apenas o sentido; (2) considerando o sentido + peso; (3) considerando valores da média da melhor testemunha; (4) considerando valores da média da melhor testemunha + peso; (5) considerando valor de uma cultivar alvo para a região e (6) considerando valor de uma cultivar alvo para a região + peso. Os cenários foram comparados por meio do ganho de seleção (GS), do ganho de seleção ajustado (GSA) e do índice de coincidência (IC). As análises demonstraram que o cenário um é o que consegue maior ganho genético total quando consideradas todas as características simultaneamente, no entanto, o cenário cinco é o que consegue maiores ganhos de seleção para produção. Tendo em vista que não foram obtidos ganhos expressivos para ciclo, que a altura de planta da população selecionada já atende ao ideótipo requerido e que o incremento da produção é o maior objetivo de um programa de melhoramento, para este conjunto de dados, a melhor construção do ideótipo foi a utilizada no cenário cinco. Portanto, os genótipos selecionados foram: MDC1312S7, MDC139S3, MDC25101S2, MDC110S2, MDC292S7, MDC137S3, MDC141S5, MDC241S5, MDC15S5, MDC246S3, MDC13S2, MDC241S6, MDC82S2, MDC248S1, MDC141S6, MDC141S2, MDC81S5, MDC91S4 e MDC255S1.

Keyword(s): *Glycine max* (L.) Merr.; *Glycine max* (L.) Merr.; Índice multi característico; MGIDI; Análise espacial; Modelos mistos; Delineamento aumentado

ABSTRACT

RODRIGUES, Fernanda Cupertino, D.Sc., Federal University of Viçosa, June 2024. **Selection strategies of soybean lines regarding statistical modeling and ideotype definition.** Advisor: Felipe Lopes da Silva. Co- advisers: Luiz Alexandre Peternelli and Maicon Nardino

In an breeding program, the choice of statistical models that provide accurate genotypic values is extremely important. However, indices that make use of various characteristics of agronomic interest for the selection of superior genotypes are also an important tool for obtaining an increase in such characteristics. This work was divided into two chapters and aimed to: (1) Compare statistical models, associated with the aforementioned analysis strategies, which favor greater efficiency in predicting genetic values of lines for agronomic characteristics of interest and (2) Compare different scenarios for construction ideotype using MGIDI to select superior soybean lines. For tabto, Cultivation Value and Use (VCU) trials were implemented in Capinópolis in the agricultural years 2020/21 and 2021/22 and in Madre de Deus in the agricultural years 20129/20 and 2021/22. The characteristics evaluated for the trials conducted in Capinópolis were: productivity (g), plant height (cm) and cycle (days), while in Madre de Deus the characteristic evaluated was productivity (g). In the first chapter, three models were tested: (1) Augmented Block Design (DBA); (2) Two-dimensional autoregressive Spatial Analysis (AR1 x AR1) and (3) Two-dimensional autoregressive Spatial Analysis with block addition. The variance components were obtained using the Restricted Maximum Likelihood (REML) and the genotypic values were predicted using Best Linear Unbased Predidction (BLUP). Each environment was analyzed separately for each variable. To compare the three models, the heritability proposed by Cullis (h^2g), Accuracy based on the heritability of Cullis ($r\hat{g}g$), Akaike Information Criterion (AIC) and Bayesian Information Criteria (BIC) were used. It was concluded that the models that best fit are those that consider the spatial correlation between plots. For three environments, the use of spatial analysis without the addition of blocks is preferred and for one environment, the model that covers spatial analysis with the addition of blocks is preferable. With the genotype values obtained through the mixed models, the MGIDI was applied in different scenarios to verify the best strategy for defining the ideotype. The ideotype was designed in six different scenarios: (1) considering only the meaning; (2) considering meaning + weight; (3) considering average values of the best witness; (4) considering average values of the best witness + weight; (5) considering the value of a target cultivar for the region and (6) considering the value

of a target cultivar for the region + weight. The scenarios were compared using the selection gain (GS), the adjusted selection gain (GSA) and the coincidence index (CI). The analyzes demonstrated that scenario one is the one that achieves the greatest total genetic gain when considering all characteristics simultaneously, however, scenario five is the one that achieves the greatest selection gains for productivity. Considering that no significant gains were obtained for the cycle, that the plant height of the selected population already meets the required ideotype and that increasing production is the main objective of a breeding program, for this set of data, the best construction of the ideotype was used in scenario five. Therefore, the selected genotypes were: MDC1312S7, MDC139S3, MDC25101S2, MDC110S2, MDC292S7, MDC137S3, MDC141S5, MDC241S5, MDC15S5, MDC246S3, MDC13S2, MDC241S6, MDC82S2, MDC248S1, 41S6, MDC141S2, MDC81S5, MDC91S4 and MDC255S1.

Keyword(s): Glycine max (L.) Merr.; Multi-characteristic index; MGIDI; Spatial analysis; Mixed models; Augmented delineation

SUMÁRIO

INTRODUÇÃO GERAL	12
REFERÊNCIAS	15
CAPÍTULO 1	17
RESUMO.....	19
ABSTRACT	20
1. INTRODUÇÃO	21
2. MATERIAIS E MÉTODOS.....	22
2.1 Experimentos	22
2.2 Análises Estatísticas.....	27
2.2.1 Modelos Estatísticos	27
2.2.2 Delineamento em Blocos Aumentados	28
2.2.3 Análise Espacial e Análise Espacial com efeito de Bloco.....	28
2.3 Teste estatístico e comparação entre os modelos.....	29
2.4 Software	31
3. RESULTADOS E DISCUSSÃO	31
3.1 Ensaio de Capinópolis	31
3.1.1 Safra 2020/21.....	31
3.1.2 Safra 2021/22.....	40
3.2 Ensaio de Madre de Deus.....	47
3.2.1 Safra 2019/20.....	47
3.2.2 Safra 2020/21.....	49
4. CONCLUSÕES	51
REFERÊNCIAS	53
CAPÍTULO 2	56
RESUMO.....	58
ABSTRACT	59
1. INTRODUÇÃO	60
2. MATERIAIS E MÉTODOS.....	61
2.1 Experimento.....	61
2.2 Análises estatísticas	62
2.2.1 Máxima Verossimilhança Restrita/Melhor Predição Linear não Viesada (REML/BLUP).....	62
2.2.2 Índice de Distância Genótipo-Ideótipo Multicaracterística (MGIDI)	63

3. RESULTADOS E DISCUSSÃO	68
3.1 Análise de Fatores.....	68
3.2 Seleção por MGIDI.....	69
4. CONCLUSÕES	80
REFERÊNCIAS	81
CONCLUSÃO GERAL	84

INTRODUÇÃO GERAL

A soja (*Glycine max* (L.) Merrill) é uma espécie autógama originária da Ásia, mais precisamente da região da Manchúria, na Índia. Atualmente, o Brasil é o maior produtor do mundo, neste ano, na safra 23/2024, o Brasil produziu mais de 147 milhões de toneladas e atingiu uma produção de 3.205 kg/ha (CONAB, 2024).

Dada a importância da cultura da soja e, para atender as exigências do mercado, os melhoristas estão sempre em busca de cultivares superiores para serem lançadas nas cinco macrorregiões sojícolas do Brasil. No entanto, a seleção incorreta de tais linhagens pode resultar na ineficiência de um programa de melhoramento (BRENNAN, 1988). Em virtude disso, a escolha de modelos estatísticos que forneçam valores genéticos acurados das linhagens são de extrema importância para o sucesso de um programa de melhoramento. Dois destes modelos estatísticos que são largamente utilizadas para o ranqueamento de genótipos são aqueles oriundos do Delineamento em Blocos Aumentados e os que consideram técnicas de Análise Espacial.

O Delineamento em Blocos Aumentados (DBA), proposto em 1956 por Walter T. Federer, é definido como um delineamento padrão, como o delineamento em blocos, por exemplo, que recebe tratamentos adicionais (FEDERER, 1956; 1958; 1961a; 1961b; FEDERER & RAGHAVARAO, 1975). Assim, o DBA possui dois conjuntos de tratamentos: tratamentos comuns/testemunhas, repetidos r vezes e tratamentos adicionais/regulares, que são responsáveis pelo aumento dos blocos (FEDERER, 1956). Durante duas décadas após a sua proposição, os delineamentos aumentados passaram por um certo esquecimento (DUARTE, 2001). Contudo, nos últimos anos, sua utilização tem crescido significativamente em programas de melhoramento genético vegetal de diversas culturas como a soja (PUTRI *et al.*, 2020; AZEVEDO *et al.*, 2020; PERTIWI *et al.*, 2022), feijão (SABA *et al.*, 2017; BASAVARAJA *et al.*, 2022, Ullah *et al.*, 2024), cana-de-açúcar (SANGHERA *et al.*, 2019; CARNEIRO *et al.*, 2020; ISLAM *et al.*, 2021), trigo (WOYANN *et al.*, 2019; GAHLAUT *et al.*, 2022), milho (JIMÉNEZ-GALINDO, *et al.*, 2019; SALUCI *et al.*, 2021) entre outras.

Resende & Sturion (2003) mencionam que, na análise convencional de experimentos de campo, espera-se que as observações coletadas em posições adjacentes (em plantas ou parcelas vizinhas) não possuam correlação, ou seja, espera-se que os erros sejam independentes. Todavia, apesar do exposto, Eisenberg *et al.* (1996) deixam claro que, embora os pesquisadores possam estar atentos aos cuidados necessários para o

planejamento de um experimento, variáveis imprevisíveis podem afetar o ensaio, fazendo com que os princípios básicos e a não correlação espacial entre plantas/parcelas vizinhas não sejam atendidos. Tais variáveis podem influenciar as análises, fazendo com que estas não sejam fidedignas e os efeitos dos tratamentos sejam tomados com um baixo grau de confiança. Para tanto, a Análise Espacial tem sido sugerida no melhoramento de plantas, a fim de exercer maior controle local e aumentar a precisão experimental (SILVA *et al.*, 2020). Esta considera a correlação espacial que pode ocorrer entre plantas/parcelas e particionam o resíduo (\mathbf{e}) em duas partes: efeito de tendência (ξ) e resíduo aleatório não correlacionado (η), ou seja, $\mathbf{e} = \xi + \eta$ (RESENDE & SATURION, 2003). A estratégia da Análise Espacial também tem sido largamente utilizada em programas de melhoramento de plantas (ELIAS *et al.*, 2018; RODRÍGUEZ-ÁLVAREZ, 2018; BERNARDELI, 2020; SILVA, 2020; GOGEL *et al.*, 2023; JANG *et al.*, 2023; TADESE *et al.*, 2023; DINSA & TEDESE *et al.*, 2024).

O incremento da produção é o principal objetivo de um programa de melhoramento, no entanto, é necessário que o material genético selecionado reúna, simultaneamente, uma série de atributos favoráveis que lhe confira rendimento comparativamente mais elevado que satisfaça as exigências do mercado. Para tanto, existem algumas ferramentas estatísticas que aumentam a chance de êxito no processo de seleção de genótipos superiores, dentre elas estão os índices de seleção (CRUZ *et al.*, 2012). Alguns desses índices permitem que a seleção seja realizada com base em várias características ao mesmo tempo. Tais índices criam valores numéricos, que funcionam como um caráter adicional a análise, resultante da combinação de determinados caracteres escolhidos pelo melhorista, sobre os quais deseja exercer a seleção simultânea (GRANATE *et al.*, 2002).

Dentre os índices disponíveis na literatura está o MGIDI (*Multi-trait Genotype–Ideotype Distance Index*), proposto por Olivoto & Nardino (2021). Quando comparados com outros índices clássicos, como o Índice Smith-Hazel (SMITH, 1936; HAZEL, 1943), o MGIDI possui como características vantajosas: i) capacidade de contornar problemas de multicolineariedade; ii) não necessita da atribuição de pesos econômicos; iii) a correlação genética é considerada, permitindo preservar as relações entre as características e fornecendo um significado biológico aos ideótipos (ROCHA *et al.*, 2017; ROCHA *et al.*, 2019; OLIVOTO & NARDINO, 2021; FERREIRA, 2020; NASCIMENTO, 2020; VOLPATO *et al.*, 2021). As análises estatísticas do índice MGIDI são realizadas em quatro principais etapas: i) na primeira, o planejamento de um

ideótipo é realizado; ii) posteriormente, a análise redimensiona os valores das características de 0-100; iii) nesta etapa, uma análise exploratória de fatores é utilizada para explicar a estrutura da correlação e a dimensionalidade da redução dos dados; e iv) por último, são calculadas as distâncias entre genótipos e ideótipos (OLIVOTO & NARDINO, 2021). Desde que o MGIDI foi proposto, programas de melhoramento de diferentes culturas passaram a utilizá-lo como estratégia de seleção de genótipos superiores com base em várias características (BENAKANNAHALLI *et al.*, 2021; POUR-ABOUGHADAREH *et al.*, 2021; POUR-ABOUGHADAREH & POCZAI, 2021; OLIVOTO *et al.*, 2021; DALAROSA, 2021; SOARES, 2022; MOREIRA, 2022; TOSOTT *et al.*, 2023; DUARTE, 2023; COSTA *et al.*, 2023; KLEIN *et al.*, 2023; PALLAVI *et al.*, 2014).

Dado o exposto, este trabalho foi dividido em dois capítulos. O primeiro objetivou comparar modelos estatísticos que forneçam maior eficiência na predição de valores genéticos de linhagens para características agronômicas de interesse do programa de melhoramento de soja. Por conseguinte, o segundo teve como objetivo comparar diferentes cenários para construção de um ideótipo com o uso do MGIDI para a seleção de linhagens superiores de soja.

REFERÊNCIAS

- AZEVEDO, Cleber Vinicius Giaretta *et al.* Genetic parameters of soybean populations obtained from crosses between grain and food genotypes. **Acta Scientiarum**. Agronomy, v. 43, 2020.
- BERNARDELI, Arthur *et al.* Modeling spatial trends and enhancing genetic selection: An approach to soybean seed composition breeding. **Crop Science**, v. 61, n. 2, p. 976-988, 2021
- BRENNAN, J.P. (1988). An economic investigation of wheat breeding programmes. **Agricultural Economics Bulletin** 35, Department of Agricultural Economics and Business Management, **University of New England, Armidale, NSW, Australia**.
- CULLIS, B.R.; GLEESON, A.C. Spatial analysis of field experiments – an extension at two dimensions. **Biometrics**, v. 47, p. 1449 – 1460, 1991.
- DUARTE, João Batista; VENCOVSKY, Roland; DIAS, Carlos Tadeu dos Santos. Estimadores de componentes de variância em delineamento de blocos aumentados com tratamentos novos de uma ou mais populações. **Pesquisa Agropecuária Brasileira**, v. 36, p. 1155-1167, 2001.
- CARNEIRO, Monalisa Sampaio *et al.* RB985476-a sugarcane cultivar with high agro-industrial yield and disease resistance. **Crop Breeding and Applied Biotechnology**, v. 20, 2020.
- CULLIS, B.R.; SMITH, A.B.; COOMBES, N.E. 2006. On the design of early generation variety trials with correlated data. **J. Agric. Biol. Environ. Stat.** 11(4): 381–393. doi: 10.1198/108571106X154443.
- EISENBERG, B.E.; GAUCH, H.G.; ZOBEL, R.W.; KILIAN, W. Spatial analysis of field experiments: fertilizer experiments with wheat (*Triticum aestivum*) and tea (*Camellia sinensis*). In: KANG, M.S.; GAUCH, H.G. (Ed.) Genotype by environment interaction. **Boca Raton: CRC Press**, 1996. Cap. 14, p. 373-404.
- FEDERER, W.T. Augmented (or hoonuiaku) designs. **Hawaiian Planter's Records**, v. 55, p. 191 – 208, 1956.
- FEDERER, W.T. Augmented designs. **Biometrics**, v. 14, p. 134 (Abstract 469), Mar 1958.
- GAHLAUT, Vijay *et al.* Dataset of nine agronomic traits in bread wheat phenotyped under irrigated and rain-fed environments. **Data in Brief**, v. 41, p. 107933, 2022.
- HAZEL, L.N. The genetic basis for constructing selection indexes. **Genetics**, v.28, p.476-490, 1943.
- HENDERSON, C.R. 1975. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics** 31(2): 423. doi: 10.2307/2529430.

ISLAM, Md S. *et al.* Experimental evaluation of genomic selection prediction for rust resistance in sugarcane. **The Plant Genome**, v. 14, n. 3, p. e20148, 2021.

JIMÉNEZ-GALINDO, José Cruz *et al.* Mapping of resistance to corn borers in a MAGIC population of maize. **BMC plant biology**, v. 19, n. 1, p. 1-17, 2019.

SALUCI, Júlio Cesar Gradice *et al.* Potential of Popcorn S4 lines for resistance to southern corn leaf blight. **Functional Plant Breeding Journal**, v. 2, n. 2, 2020.

RESENDE, M. D. V.; STURION, J. A. Análise estatística espacial de experimentos via modelos mistos individuais com erros modelados por processos ARIMA em duas dimensões. **Revista de matemática e estatística**, v. 21, p. 7-33, 2003.

OLIVOTO, Tiago *et al.* MGIDI: a powerful tool to analyze plant multivariate data. **Plant Methods**, v. 18, n. 1, p. 121, 2022.

PALLAVI, M. *et al.* Multi trait genotype- ideotype distance index (MGIDI) for early seedling vigour and yield related traits to identify elite lines in rice (*Oryza sativa* L.). **Electronic Journal of Plant Breeding**. Vol 15(1) : 120 – 131

POUR-ABOUGHADAREH, Alizera *et al.* Identification of salt-tolerant barley genotypes using multiple-traits index and yield performance at the early growth and maturity stages. **Bulletin of the National Research Centre** (2021) 45:117

RESENDE, M.D.V., SILVA, F.F.; AZEVEDO, C.F. **Estatística matemática**. 2014.

CAPÍTULO 1

**DELINEAMENTO DE BLOCOS AUMENTADOS *versus* ANÁLISE ESPACIAL
NA PREDIÇÃO DE VALORES GENÉTICOS DE LINHAGENS DE SOJA**

RESUMO

RODRIGUES, Fernanda Cupertino, D.Sc., Universidade Federal de Viçosa, junho de 2024. **Delineamento de blocos aumentados versus análise espacial na predição de valores genéticos de linhagens de soja.** Orientador: Felipe Lopes da Silva. Co-orientadores: Luiz Alexandre Peternelli e Maicon Nardino

O programas de melhoramento de soja almejam cultivares mais produtivas e adaptadas às regiões alvo. Em busca disso, muitos ensaios são realizados e as análises estatísticas são uma etapa de grande importância para a avaliação e seleção de cultivares elites. Com o objetivo de definir os modelos estatísticos que mais se ajustam à ensaios de Valor de Cultivo e Uso (VCU), foram implantados em Capinópolis nos anos agrícola de 2020/21 e 2021/22 e em Madre de Deus nos anos agrícolas de 20129/20 e 2021/22. As características avaliadas para os ensaios implantados em Capinópolis foram: produção (g), altura de planta (cm) e ciclo (dias) e para Madre de Deus a característica avaliada foi produção. Nas análises estatísticas três modelos foram testados. O primeiro modelo corresponde ao Delineamento em Blocos Aumentados (DBA); o segundo corresponde à Análise Espacial autorregressiva de duas dimensões (AR1 x AR1); e o terceiro corresponde à Análise Espacial autorregressiva de duas dimensões com adição de bloco. As análises foram realizadas no *software R* utilizando o pacote *asreml*. Os componentes de variância foram obtidos pela metodologia de modelos mistos *Restricted Maximum Likelihood* (REML) e os valores genotípicos foram preditos utilizando *Best Linear Unbased Prediction* (BLUP). Analisou-se cada ambiente separadamente para cada variável. Para testar os modeos realizouse realizado *Likelihood Ratio Test* (LRT) e para comparar os três modelos foram utilizadas a herdabilidade proposta por *Cullis* (h^2_g), a Acurácia com base na herdabilidade de *Cullis* (r_{gg}), *Akaike Information Criterion* (AIC) e *Bayesian Information Criteria* (BIC). Concluiu-se que os modelos que melhor se ajustam são os que consideram a correlação espacial entre as parcelas. Para três ambientes é preferencial a utilização da análise espacial sem a adição de blocos e para um ambiente é preferível o modelo que abrange a análise espacial com a adição de blocos.

Keyword(s): Glycine max (L.) Merr.; Spatial analysis; Mixed models; Augmented delineation

ABSTRACT

RODRIGUES, Fernanda Cupertino, D.Sc., Federal University of Viçosa, June 2024. **Augmented block design versus spatial analysis in predicting genetic values of soybean lines.** Advisor: Felipe Lopes da Silva. Co-advisors: Luiz Alexandre Peternelli and Maicon Nardino.

Soybean breeding programs aim for more productive cultivars adapted to the target regions. In pursuit of this, many tests are carried out and statistical analyzes are a very important step for the evaluation and selection of elite cultivars. With the aim of defining the statistical models that best fit the Value of Cultivation and Use (VCU) trials implemented in Capinópolis in the agricultural years 2020/21 and 2021/22 and in Madre de Deus in the agricultural years 20129/20 and 2021/22, three models were tested. The first model corresponds to the Augmented Block Design (DBA); the second corresponds to the two-dimensional autoregressive Spatial Analysis (AR1 x AR1); and the third corresponds to two-dimensional autoregressive Spatial Analysis with block addition. The characteristics evaluated for the implemented trials conducted in Capinópolis were: productivity, plant height and cycle. The characteristic evaluated for the tests in Madre de Deus was productivity. The analyzes were performed in R software using the *asreml* package. The variance components were obtained using the Restricted Maximum Likelihood (REML) and the genotypic values were predicted using Best Linear Unbased Prediction (BLUP). Each environment was analyzed separately for each variable. To test the models, the Likelihood Ratio Test (LRT) was performed and to compare the three models, the heritability proposed by Cullis (h^2_g), the Accuracy based on Cullis heritability (r_{gg}), Akaike Information Criterion (AIC) and Bayesian Information Criteria (BIC) were used. It was concluded that the models that best fit are those that consider the spatial correlation between plots. For three environments, the use of spatial analysis without the addition of blocks is preferred and for one environment, the model that covers spatial analysis with the addition of blocks is preferable.

1. INTRODUÇÃO

O objetivo do melhoramento genético da soja é selecionar linhagens superiores mais produtivas e adaptadas para as regiões alvo. No entanto, a seleção incorreta de tais linhagens pode resultar na ineficiência de um programa de melhoramento (BRENNAN, 1988). Em virtude disso, a escolha de modelos estatísticos que forneçam valores genéticos acurados das linhagens são de extrema importância para o sucesso de um programa de melhoramento. Dois destes modelos estatísticos que são largamente utilizadas para o ranqueamento de genótipos são aqueles oriundos do Delineamento em Blocos Aumentados e os que consideram técnicas de Análise Espacial.

O Delineamento em Blocos Aumentados (DBA), proposto em 1956 por Walter T. Federer, é definido como um delineamento padrão, como o delineamento em blocos, por exemplo, que recebe tratamentos adicionais (FEDERER, 1956; 1958; 1961a; 1961b; FEDERER; RAGHAVARAO, 1975). Assim, o DBA possui dois conjuntos de tratamentos: tratamentos comuns/testemunhas, repetidos r vezes e tratamentos adicionais/regulares, que são responsáveis pelo aumento dos blocos (FEDERER, 1956). Durante duas décadas após a sua proposição, os delineamentos aumentados passaram por um certo esquecimento (DUARTE, 2001). Contudo, nos últimos anos, sua utilização tem crescido significativamente em programas de melhoramento genético vegetal de diversas culturas como a soja (PUTRI *et al.*, 2020; AZEVEDO *et al.*, 2020; PERTIWI *et al.*, 2022), feijão (SABA *et al.*, 2017; BASAVARAJA *et al.*, 2022, Ullah *et al.*, 2024), cana-de-açúcar (SANGHERA *et al.*, 2019; CARNEIRO *et al.*, 2020; ISLAM *et al.*, 2021), trigo (WOYANN *et al.*, 2019; GAHLAUT *et al.*, 2022), milho (JIMÉNEZ-GALINDO, *et al.*, 2019; SALUCI *et al.*, 2021) entre outras.

Resende & Sturion (2003) mencionam que na análise convencional de experimentos de campo, espera-se que as observações coletadas em posições adjacentes (em plantas ou parcelas vizinhas) não possuam correlação, ou seja, espera-se que os erros sejam independentes. Todavia, apesar do exposto, Eisenberg *et al.* (1996), deixam claro que, embora os pesquisadores possam estar atentos aos cuidados necessários para o planejamento de um experimento, variáveis imprevisíveis podem afetar o ensaio, fazendo com que os princípios básicos e a não correlação espacial entre plantas/parcelas vizinhas não sejam atendidos. Tais variáveis podem influenciar as análises, fazendo com que estas não sejam fidedignas e os efeitos dos tratamentos sejam tomados com um baixo grau de confiança. Para tanto, a Análise Espacial têm sido sugerida no melhoramento de plantas,

a fim de exercer maior controle local e aumentar a precisão experimental (SILVA *et al.*, 2020). Esta considera a correlação espacial que pode ocorrer entre plantas/parcelas e particionam o resíduo (\mathbf{e}) em duas partes: efeito de tendência (ξ) e resíduo aleatório não correlacionado (η), ou seja, $\mathbf{e} = \xi + \eta$ (RESENDE & SATURION, 2003). A estratégia da Análise Espacial também tem sido largamente utilizada em programas de melhoramento de plantas (ELIAS *et al.*, 2018; RODRÍGUEZ-ÁLVAREZ, 2018; BERNARDELI, 2020; SILVA, 2020; GOGEL *et al.*, 2023, JANG *et al.*, 2023; TADESE *et al.*, 2023; DINSA & TEDESE *et al.*, 2024)

Tendo em vista que tanto o Delineamento em Blocos Aumentados quanto a Análise Espacial são amplamente utilizados em programas de melhoramento de plantas, o objetivo deste trabalho foi comparar modelos estatísticos, associados às estratégias de análises citadas, que favoreçam a maior eficiência na predição de valores genéticos de linhagens para características agrônômicas de interesse de programas de melhoramento de soja.

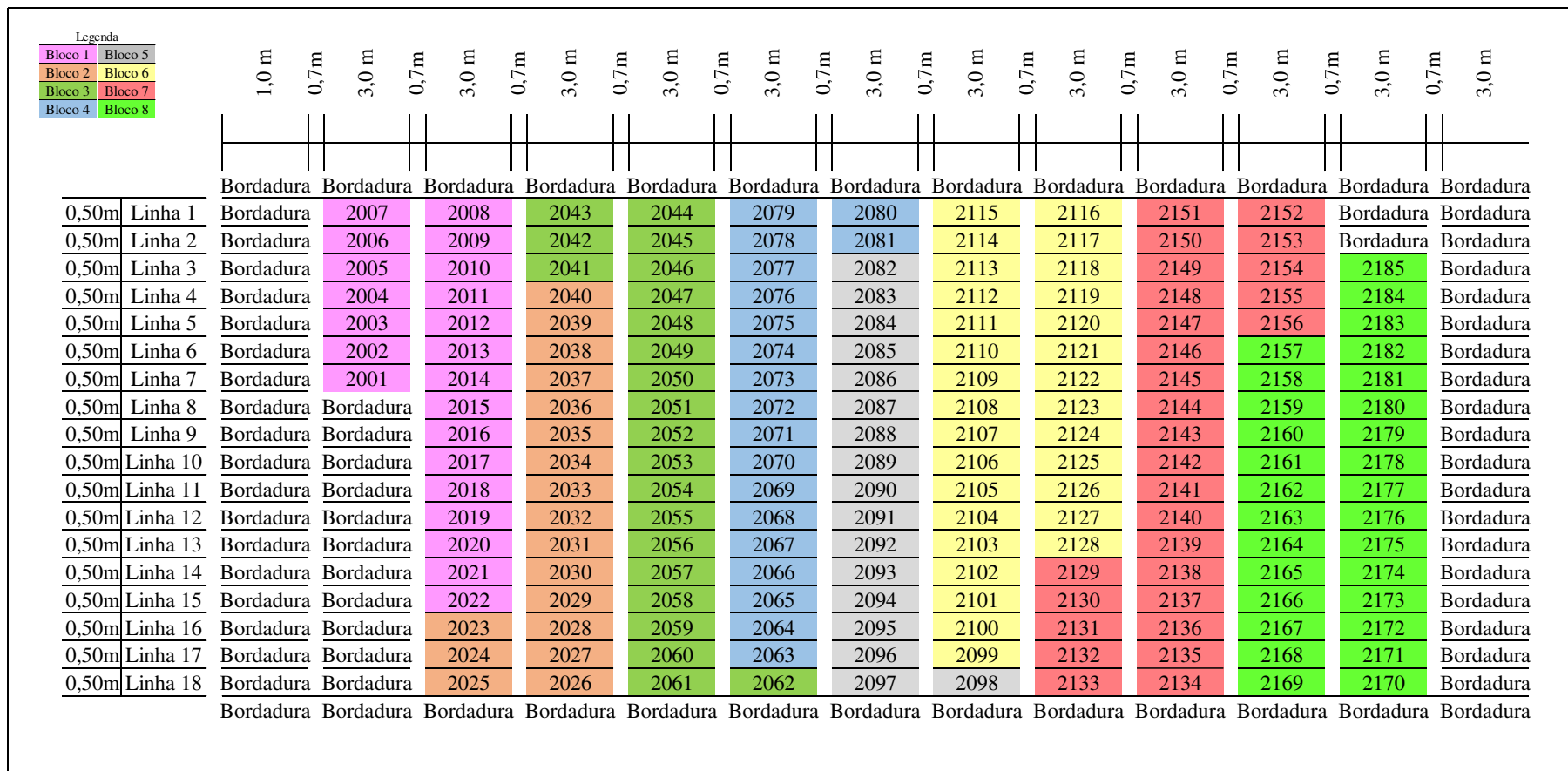
2. MATERIAIS E MÉTODOS

2.1 Experimentos

Com o objetivo de testar diferentes ensaios e condições ambientais, quatro experimentos foram conduzidos para o presente estudo, sendo implantados em localidades no Estado de Minas Gerais: Capinópolis (18°40'48"S, 49°33'58"W, 530 metros de altitude) e Madre de Deus de Minas (21° 29' 2"S, 44°19'58"W, 953 metros de altitude). Todos os ensaios foram instalados em Delineamento de Blocos Aumentados e plantados com densidade de semeadura de 12 sementes por metro.

Os ensaios em Capinópolis foram conduzidos durante os anos agrícolas de: (1) 2020/21, sendo constituído de 246 genótipos regulares e seis testemunhas, distribuídos em parcelas experimentais de duas linhas de 2,5 m cada. Os genótipos regulares foram distribuídos em oito blocos e as testemunhas foram alocadas em todos os blocos, totalizando 294 parcelas (Figura 1) e (2) 2021/22, sendo constituído de 128 genótipos regulares e seis testemunhas distribuídos em parcelas de uma linha de 3 m. Os genótipos regulares foram distribuídos em oito blocos e as testemunhas foram alocadas em todos os blocos, totalizando 176 parcelas (Figura 2).

Figura 2 – Croqui da área experimental em Cainópolis (MG) na safra de 2021/22
Capinópolis Safra 2021/22



Fonte: Elaboração própria.

Os ensaios em Madre de Deus de Minas foram instalados nos anos agrícolas de: (1) 2019/20, sendo constituído de 90 genótipos regulares e 6 testemunhas, distribuídos parcelas de duas linhas de 2,5 m (Figura 3). Os genótipos regulares foram distribuídos em cinco blocos e as testemunhas foram alocadas em todos os blocos, totalizando 120 parcelas e (2) 2021/22, sendo constituído de 129 genótipos regulares e seis testemunhas, distribuídos em parcelas foi constituída de duas linhas de 3,0 m cada, os genótipos regulares foram distribuídos em oito blocos e as testemunhas alocadas em todos os locos, totalizando 177 parcelas (Figura 4).

Figura 3 – Croqui da área experimental em Madre de Deus de Minas (MG) na safra de 2020/21.

Legenda															
Bloco 1	Bloco 5	1,5 m	0,7m	2,5 m	0,7m	2,5 m	0,7m	2,5 m	0,7m	2,5 m	0,7m	2,5 m	0,7m	2,5 m	0,7m
Bloco 2	Bloco 6	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura
Bloco 3	Bloco 7	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m	0,50m
Bloco 4	Bloco 8	Linha 1	Linha 2	Linha 3	Linha 4	Linha 5	Linha 6	Linha 7	Linha 8	Linha 9	Linha 10	Linha 11	Linha 12	Linha 13	Linha 14
		1001	1048	1049	1096	1097	1144	1145	1001	1048	1049	1096	1097	1144	1145
		1002	1047	1050	1095	1098	1143	1146	1002	1047	1050	1095	1098	1143	1146
		1003	1046	1051	1094	1099	1142	1147	1003	1046	1051	1094	1099	1142	1147
		1004	1045	1052	1093	1100	1141	1148	1004	1045	1052	1093	1100	1141	1148
		1005	1044	1053	1092	1101	1140	1149	1005	1044	1053	1092	1101	1140	1149
		1006	1043	1054	1091	1102	1139	1150	1006	1043	1054	1091	1102	1139	1150
		1007	1042	1055	1090	1103	1138	1151	1007	1042	1055	1090	1103	1138	1151
		1008	1041	1056	1089	1104	1137	1152	1008	1041	1056	1089	1104	1137	1152
		1009	1040	1057	1088	1105	1136	1153	1009	1040	1057	1088	1105	1136	1153
		1010	1039	1058	1087	1106	1135	1154	1010	1039	1058	1087	1106	1135	1154
		1011	1038	1059	1086	1107	1134	1155	1011	1038	1059	1086	1107	1134	1155
		1012	1037	1060	1085	1108	1133	1156	1012	1037	1060	1085	1108	1133	1156
		1013	1036	1061	1084	1109	1132	1157	1013	1036	1061	1084	1109	1132	1157
		1014	1035	1062	1083	1110	1131	1158	1014	1035	1062	1083	1110	1131	1158
		1015	1034	1063	1082	1111	1130	1159	1015	1034	1063	1082	1111	1130	1159
		1016	1033	1064	1081	1112	1129	1160	1016	1033	1064	1081	1112	1129	1160
		1017	1032	1065	1080	1113	1128	1161	1017	1032	1065	1080	1113	1128	1161
		1018	1031	1066	1079	1114	1127	1162	1018	1031	1066	1079	1114	1127	1162
		1019	1030	1067	1078	1115	1126	1163	1019	1030	1067	1078	1115	1126	1163
		1020	1029	1068	1077	1116	1125	1164	1020	1029	1068	1077	1116	1125	1164
		1021	1028	1069	1076	1117	1124	1165	1021	1028	1069	1076	1117	1124	1165
		1022	1027	1070	1075	1118	1123	1166	1022	1027	1070	1075	1118	1123	1166
		1023	1026	1071	1074	1119	1122	1167	1023	1026	1071	1074	1119	1122	1167
		1024	1025	1072	1073	1120	1121	1168	1024	1025	1072	1073	1120	1121	1168
		1024	1025	1072	1073	1120	1121	1168	1024	1025	1072	1073	1120	1121	1168

Fonte: Elaboração própria.

Figura 4 – Croqui da área experimental em Madre de Deus de Minas (MG) na safra de 2020/21.

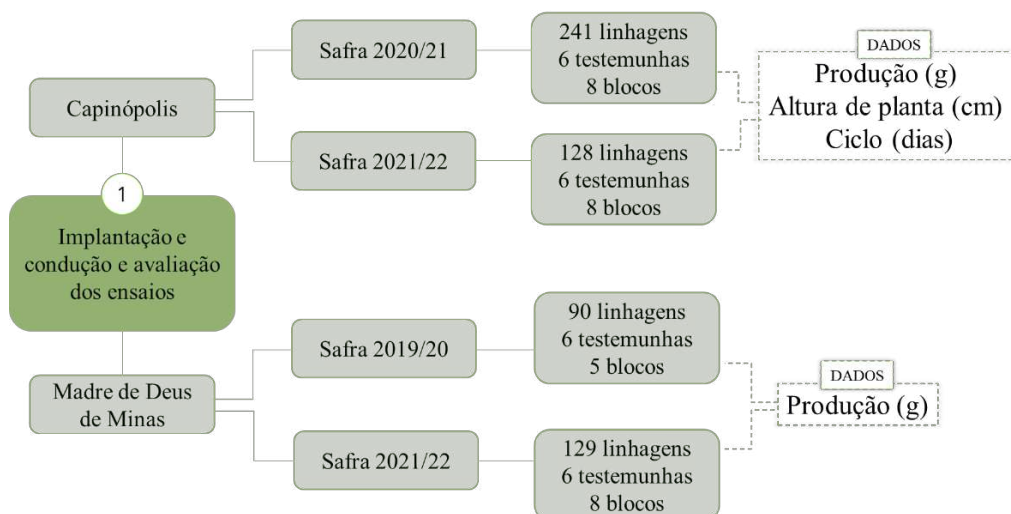
Legenda																							
Bloco 1	Bloco 5	1,5 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	3 m	0,7 m	1,5 m	
Bloco 2	Bloco 6																						
Bloco 3	Bloco 7																						
Bloco 4	Bloco 8																						
		Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura
0,50m	Linha 1	Bordadura	2001	2044	2045	2088	2089	2132	2133	2176	2177												
0,50m	Linha 2	Bordadura	2002	2043	2046	2087	2090	2131	2134	2175	2178												
0,50m	Linha 3	Bordadura	2003	2042	2047	2086	2091	2130	2135	2174	2179												
0,50m	Linha 4	Bordadura	2004	2041	2048	2085	2092	2129	2136	2173	2180												
0,50m	Linha 5	Bordadura	2005	2040	2049	2084	2093	2128	2137	2172	2181												
0,50m	Linha 6	Bordadura	2006	2039	2050	2083	2094	2127	2138	2171	2182												
0,50m	Linha 7	Bordadura	2007	2038	2051	2082	2095	2126	2139	2170	2183												
0,50m	Linha 8	Bordadura	2008	2037	2052	2081	2096	2125	2140	2169	2184												
0,50m	Linha 9	Bordadura	2009	2036	2053	2080	2097	2124	2141	2168	2185												
0,50m	Linha 10	Bordadura	2010	2035	2054	2079	2098	2123	2142	2167	Bordadura												
0,50m	Linha 11	Bordadura	2011	2034	2055	2078	2099	2122	2143	2166	Bordadura												
0,50m	Linha 12	Bordadura	2012	2033	2056	2077	2100	2121	2144	2165	Bordadura												
0,50m	Linha 13	Bordadura	2013	2032	2057	2076	2101	2120	2145	2164	Bordadura												
0,50m	Linha 14	Bordadura	2014	2031	2058	2075	2102	2119	2146	2163	Bordadura												
0,50m	Linha 15	Bordadura	2015	2030	2059	2074	2103	2118	2147	2162	Bordadura												
0,50m	Linha 16	Bordadura	2016	2029	2060	2073	2104	2117	2148	2161	Bordadura												
0,50m	Linha 17	Bordadura	2017	2028	2061	2072	2105	2116	2149	2160	Bordadura												
0,50m	Linha 18	Bordadura	2018	2027	2062	2071	2106	2115	2150	2159	Bordadura												
0,50m	Linha 19	Bordadura	2019	2026	2063	2070	2107	2114	2151	2158	Bordadura												
0,50m	Linha 20	Bordadura	2020	2025	2064	2069	2108	2113	2152	2157	Bordadura												
0,50m	Linha 21	Bordadura	2021	2024	2065	2068	2109	2112	2153	2156	Bordadura												
0,50m	Linha 22	Bordadura	2022	2023	2066	2067	2110	2111	2154	2155	Bordadura												
		Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura	Bordadura

Fonte: Elaboração própria.

Para todos os ensaios, as testemunhas utilizadas foram: BMX Zeus IPRO (Grupo Maturidade Relativa (GMR) = 5.5), BMX Alvo RR (GMR = 5.9), BS2606 IPRO (GMR = 6.0), TMG 7063 IPRO (GMR = 6.3), DM 66i68 RSF IPRO (GMR= 6.6) e FPS Júpiter RR (GMR= 6.0).

Avaliou-se as seguintes características (Figura 5): produção de grãos por parcela, ciclo e altura da planta na maturação. As parcelas foram colhidas, trilhadas em bulk e as sementes foram pesadas (gramas) para compor o dado de produção por parcela em gramas. A altura de planta na maturação, em centímetros, foi medida quando mais de 50% das plantas se encontravam no estágio R8, conforme estabelecido por Fehr e Cavinnes (1977). Com auxílio de uma régua, a planta foi medida da base à extremidade da haste principal. Para o ciclo, estabeleceu o período em dias entre a data de semeadura e a data de maturação, quando mais de 50% das plantas na se encontravam no estágio R8. Os tratos culturais durante o ciclo foram conduzidos segundo as necessidades da cultura da soja para essa região (SEDIYAMA; SILVA; BOREM, 2022).

Figura 5 – Fluxograma dos experimentos implementados em Capinópolis (MG) e Madre de Deus de Minas (MG).



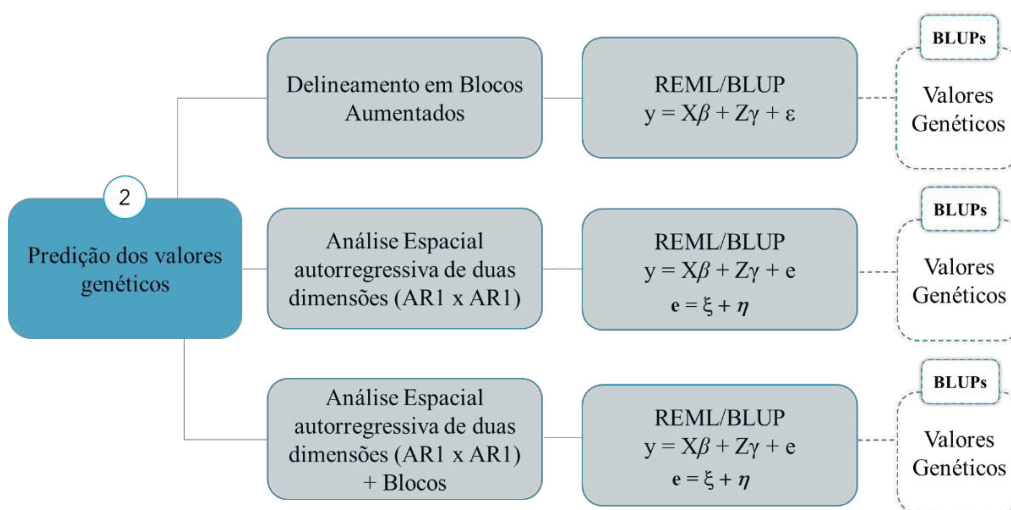
Fonte: Elaboração própria.

2.2 Análises Estatísticas

2.2.1 Modelos Estatísticos

As análises foram realizadas considerando três estratégias: modelo estatístico com manutenção da estrutura do delineamento experimental original, DBA; uso da análise espacial com modelo auto-regressivo de duas dimensões; e uso de análise espacial com modelo auto-regressivo de duas dimensões com inclusão do efeito de bloco (Figura 6).

Figura 6 – Fluxograma das análises estatísticas realizadas nos experimentos.



Fonte: Elaboração própria.

2.2.2 Delineamento em Blocos Aumentados

Na análise dos dados os componentes de variância foram obtidos pela metodologia de modelos mistos Máxima Verossimilhança Restrita – REML (PATTERSON & THOMPSON, 1971) e os valores genotípicos foram preditos utilizando Melhor Predição Linear não Viesada – BLUP (HENDERSON, 1975). As testemunhas e os blocos foram tratados como de efeitos fixos e os genótipos regulares como de efeitos aleatórios. O modelo estatístico utilizado foi o proposto por Federer (1956):

$$y = X\beta + Z\gamma + \varepsilon \text{ (modelo 1)}$$

Onde,

y é o vetor dos dados;

β é o vetor dos efeitos fixos, com a matriz de incidência X ;

γ é o vetor dos efeitos aleatórios;

ε é o vetor do erro, $\varepsilon \sim \text{NID}(0, I\sigma^2_e)$, onde σ^2_e é a variância residual.

X é a matriz de incidência dos efeitos contidos em β ;

Z é a matriz de incidência dos efeitos contidos em γ e

2.2.3 Análise Espacial e Análise Espacial com efeito de Bloco

Os componentes de variância também foram obtidos pela metodologia de modelos mistos Máxima Verossimilhança Restrita – REML (PATTERSON & THOMPSON, 1971) e os valores genotípicos foram preditos utilizando Melhor Predição Linear não Viesada – BLUP (HENDERSON, 1975). O método utilizado foi o de análises de séries temporais auto-regressivo de duas dimensões (AR1 x AR1). As análises foram realizadas de duas maneiras: auto-regressivo de duas dimensões (modelo 2) e auto-regressivo de duas dimensões com efeito de bloco (modelo 3). As testemunhas e os blocos foram tratados como de efeitos fixos e os genótipos regulares como de efeitos aleatórios. O modelo utilizado foi o geral proposto por Henderson (1984):

$$y = X\beta + Z\gamma + e \text{ (modelo 2)}$$

Onde,

y é o vetor dos dados;

β é o vetor paramétrico dos efeitos fixos (testemunhas);

X é a matriz de incidência dos efeitos contidos em β ;

γ é o vetor paramétrico dos efeitos aleatórios;

Z é a matriz de incidência dos efeitos contidos em γ e

e são os erros que compõem o vetor, onde ξ é o efeito de tendência e η é o resíduo aleatório não correlacionado).

Finalmente, o modelo 3 também pode ser representado por:

$$y = X\beta + Z\gamma + e \text{ (modelo 3)}$$

Onde,

y é o vetor dos dados;

β é o vetor paramétrico dos efeitos fixos (testemunhas);

X é a matriz de incidência dos efeitos contidos em β ;

γ é o vetor paramétrico dos efeitos aleatórios (genótipo e bloco);

Z é a matriz de incidência dos efeitos contidos em γ e

e são os erros que compõem o vetor ($e = \xi + \eta$, onde ξ é o efeito de tendência e η é o resíduo aleatório não correlacionado).

2.3 Teste estatístico e comparação entre os modelos

A significância da diferença entre os modelos foi testada via LRT (*Likelihood Ratio Test*) por meio da equação abaixo:

$$LRT = -2 (\text{Log}L - \text{Log}L_r)$$

Onde,

$\text{Log}L$ é o logaritmo da função de verossimilhança do modelo completo e

$\text{Log}L_r$ é o logaritmo da função de verossimilhança do modelo reduzido.

Para comparar os três modelos foram utilizadas a herdabilidade proposta por Cullis (h^2_g), a Acurácia com base na herdabilidade de Cullis ($r_{\hat{g}g}$), *Akaike Information Criterion* (AIC) e *Bayesian Information Criteria* (BIC).

De acordo com Cullis *et al.* (2006), a h^2_g se diferencia por não utilizar a variância total, ou seja, a variância genética somada à variância do resíduo. Cullis *et al.* (2006) defende que em experimentos com dados balanceados a herdabilidade convencional utilizando variância total ou a de Cullis utilizando apenas os valores genéticos não vão se diferir, mas para dados desbalanceados a herdabilidade que considera a variância total não reflete o ganho de seleção. Este parâmetro é calculado da seguinte maneira:

$$h^2_g = 1 - \frac{\bar{v}_{BLUP}}{2\sigma_g^2} \quad (\text{Cullis et al., 2006})$$

Onde,

h^2_g = herdabilidade de Cullis;

\bar{v}_{BLUP} = a variância média das diferenças entre pares de BLUPs refletindo os efeitos genotípicos;

σ_g^2 = variância genotípica.

A acurácia foi calculada com base na herdabilidade de Cullis da seguinte maneira:

$$r_{\hat{g}g} = \sqrt{h_g^2}$$

Onde,

h_g^2 é a herdabilidade de cada característica.

O AIC e o BIC são critérios matemáticos para avaliar o quão bem um modelo se ajusta aos dados e são utilizados para comparar diferentes modelos estatísticos. Quanto menor o valor para estes critérios, mais ajustado o modelo estará aos dados. Estes critérios são calculados da seguinte maneira:

$$AIC = -2\log(L) + 2p \quad (\text{Akaike, 1974})$$

Onde,

AIC= *Akaike Informatin Criterion*;

$\log(L)$ = é o logaritmo da máxima de verossimilhança do modelo M, i.e. $\hat{L} = p(x|\hat{\Theta}, M)$, onde $\hat{\Theta}$ são os parâmetros dos valores que maximizam a função de verossimilhança e x é o dado observado e

p = número de parâmetros estimados.

$$\text{BIC} = -2\log(\hat{L}) + k \log(n) \quad (\text{Schwarz, 1978})$$

Onde,

$\text{BIC} = \text{Bayesian Information Criteria}$;

\hat{L} = é o valor máximo de verossimilhança do modelo M , i.e. $\hat{L} = p(x|\hat{\theta}, M)$, onde $\hat{\theta}$ são os parâmetros dos valores que maximizam a função de verossimilhança e x é o dado observado;

n = o número de observações e

k = numero de parâmetros estimados pelo modelo.

2.4 Software

As análises foram realizadas com o auxílio do pacote ASReml (Gilmour et al. & Thompson, 2015) por meio do R 2023.03.1 (R Core Team, 2023).

3. RESULTADOS E DISCUSSÃO

Bernardelli *et al.* (2021) enfatizam em seu trabalho que em condições ambientais diferentes, o resultado das análises para as variáveis podem apresentar resultados diferentes, visto que as interações do genótipo podem não ser as mesmas. Em detrimento disso e, dado que os ensaios implantados em Capinópolis e Madre de Deus foram conduzidos em diferentes áreas das fazendas e em diferentes safras, as análises foram realizadas para cada variável e ambiente separadamente e os resultados discussões assim serão apresentados.

3.1 Ensaios de Capinópolis

3.1.1 Safra 2020/21

Para a variável produção não houve diferença significativa quando os modelos 1 e 2 foram comparados ($p > 0,05$), assim, não há diferença significativa entre DBA e

Análise Espacial (Tabela 1). No entanto, quando foram comparados os modelos 1 e 3 e os modelos 2 e 3 houve diferença significativa ($p < 0,001$) para ambas as comparações.

Tabela 1 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável produção para Capinópolis safra 2020/21.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Produção	Modelo 1 x Modelo 2	-1,8074	0,5000ns
	Modelo 1 x Modelo 3	11,9520	0,0009***
	Modelo 2 x Modelo 3	13,7590	0,0001***

LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Para os modelos que consideram dependência espacial, as variâncias do resíduo possuem valores mais elevados quando comparados com a análise em DBA e, tendo em vista que a análise espacial captura o erro aleatório e erro de tendência, as análises também sugerem que há dependência espacial entre as parcelas (Tabela 2), o que não é significativo quando se compara DBA com Análise Espacial. No entanto, é significativo quando comparado DBA com Análise Espacial com efeito de bloco.

Tabela 2 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável produção para a safra 2020/21 em Capinópolis.

	Modelo 1	Modelo 2	Modelo 3
σ_g^2	0,00511	9486,03000	0,01251
σ_b^2	14139,52367	-	12286,36000
σ_r^2	50470,54649	-	-
R	-	53152,94000	51030,14000
h_g^2	0,00001	0,11782	0,00003
r_{gg}	0,00263	0,34324	0,00505
AIC	3364,78264	3368,59000	3356,83059
BIC	3375,69771	3383,14342	3375,02236

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Informatin Criterion* e BIC = *Bayesian Information Criteria*.

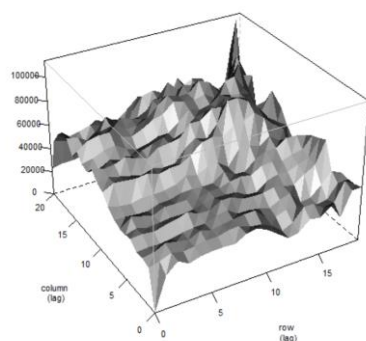
Fonte: Elaboração própria.

Por meio dos critérios de comparação calculados (Tabela 2), tanto para BIC quanto para AIC, foram obtidos menores valores para o modelo que considera efeito de blocos e dependência espacial, portanto, o modelo 3 é o que melhor se ajusta para as análises deste ensaio para esta característica. Esse mesmo resultado é facilmente observado ao analisar o variograma plotado na área do ensaio quando se comparam os modelos 2 e 3 (Figura 7). No variograma do modelo 3, em que o bloco é considerado,

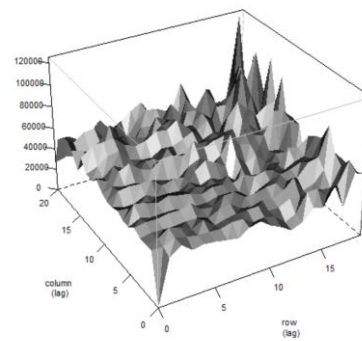
observa-se valores menores para os resíduos e menor variação em área total, ou seja, o efeito de bloco auxiliou na captura da heterogeneidade do ensaio.

Figura 7 – Variograma da variável produção para o modelo 2 (a) e modelo 3 (b) do ensaio de Capinópolis da safra 2020/21.

(a) Modelo 2



(b) Modelo 3



Fonte: Elaboração própria.

Tabela 3 – Índice de coincidência (IC) entre os genótipos selecionados para a variável produção para cada modelo estudado e para a safra 2020/21 em Capinópolis.

Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
84%	92%	89%

Fonte: Elaboração própria.

Características com herdabilidades mais baixas, como no caso da produção para este ensaio, são mais sensíveis às variações, o que valida o maior ajuste do modelo espacial com efeito de bloco para estes dados. Cullis *et al.* (1998), Bernardeli *et al.* (2019) e Ferreira (2021) em estudos realizados com as culturas da cevada, soja e tomate, respectivamente, analisaram modelos que incluem efeito de bloco na análise espacial, também concluíram que, para os dados analisados, os modelos espaciais com informação de bloco foram os que mais se ajustaram.

Tabela 4 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à produção para Capinópolis safra 2020/21.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC82S4	628,2401	MDC82S4	799,7804	MDC82S4	628,2402
2	MDC19S2	628,2401	MDC246S4	776,9801	MDC19S2	628,2402
3	MDC246S4	628,2401	MDC19S2	766,1870	MDC246S4	628,2402
4	MDC2510IS1	628,2401	MDC2510IS1	741,4175	MDC2510IS1	628,2402
5	MDC137S2	628,2401	MDC152S7	720,6691	MDC137S2	628,2401

6	MDC719S4	628,2400	MDC719S4	719,4663	MDC719S4	628,2401
7	MDC714S5	628,2400	MDC751S3	716,4258	MDC152S7	628,2401
8	MDC110S2	628,2400	MDC137S2	708,6317	MDC751S3	628,2401
9	MDC751S5	628,2400	MDC292S5	707,7388	MDC11S6	628,2401
10	MDC11S6	628,2400	MDC81S2	700,1048	MDC110S2	628,2401
11	MDC138S4	628,2400	MDC11S6	699,7989	MDC138S4	628,2401
12	MDC152S7	628,2400	MDC719S3	698,3896	MDC751S6	628,2401
13	MDC751S6	628,2400	MDC138S4	697,2339	MDC714S5	628,2401
14	MDC75S1	628,2400	MDC13S2	694,5597	MDC81S2	628,2401
15	MDC751S3	628,2400	MDC751S6	692,4794	MDC13S2	628,2401
16	MDC712S1	628,2400	MDC141S6	689,1632	MDC292S5	628,2401
17	MDC81S2	628,2400	MDC292S6	688,5609	MDC139S3	628,2401
18	MDC75S6	628,2400	MDC110S2	687,8253	MDC75S1	628,2401
19	MDC712S6	628,2400	MDC714S5	686,5853	MDC75S6	628,2401
20	MDC261S7	628,2400	MDC110S3	686,2623	MDC719S3	628,2401
21	MDC248S1	628,2400	MDC75S6	685,2335	MDC751S5	628,2401
22	MDC15S5	628,2400	MDC710S5	683,8416	MDC81S6	628,2401
23	MDC17S2	628,2400	MDC139S3	682,4491	MDC110S3	628,2401
24	MDC139S3	628,2400	MDC751S7	682,4342	MDC292S6	628,2401
25	MDC714S2	628,2400	MDC81S6	680,4309	MDC15S5	628,2401
26	MDC13S2	628,2400	MDC75S1	679,2752	MDC86S1	628,2401
27	MDC712S3	628,2400	MDC261S7	676,3332	MDC141S6	628,2401
28	MDC81S6	628,2400	MDC15S5	676,0044	MDC710S5	628,2401
29	MDC248S5	628,2400	MDC712S3	675,1934	MDC712S6	628,2401
30	MDC710S5	628,2400	MDC86S1	674,5815	MDC261S7	628,2401
31	MDC292S6	628,2400	MDC715S6	674,1109	MDC712S3	628,2401
32	MDC292S5	628,2400	MDC74S6	673,6174	MDC712S1	628,2401
33	MDC719S3	628,2400	MDC712S6	672,3279	MDC751S7	628,2401
34	MDC86S1	628,2400	MDC241S6	671,8274	MDC248S1	628,2401
35	MDC81S5	628,2400	MDC85S5	670,6600	MDC85S5	628,2401
36	MDC141S6	628,2400	MDC751S5	668,9135	MDC248S5	628,2401
37	MDC110S3	628,2400	MDC266S8	667,4942	MDC141S2	628,2401

Fonte: Elaboração própria.

Para altura de planta na maturação não houve diferença significativa ($p > 0,05$) para as comparações entre DBA e Análise espacial e DBA e Análise Espacial com adição de blocos (Tabela 5). No entanto, entre Análise Espacial e Análise Espacial com bloco houve diferença a 0,1% de significância. Tais resultados demonstram que não há correlação espacial entre as parcelas, no entanto, a blocagem é necessária para garantir maior homogeneidade ao experimento. Isso também pode ser evidenciado por meio do índice de coincidência, em que a coincidência entre os genótipos selecionados (Tabela 8) nos modelos que consideram efeito de bloco é maior (86%) quando comparado com os modelos que não consideram efeito de bloco (Tabela 7).

Tabela 5 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável altura para Capinópolis safra 2020/21.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Altura	Modelo 1 x Modelo 2	-4,8901	0,5000 ^{ns}
	Modelo 1 x Modelo 3	0,1504	0,2279 ^{ns}
	Modelo 2 x Modelo 3	6,3943	0,0057***

LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Ao analisar os critérios de comparação para os modelos adotados, também foi evidenciado que o modelo mais ajustado para a análise dessa característica neste ensaio é o modelo 1, o DBA, ou seja, para este conjunto de dados o delineamento em blocos foi eficiente em corrigir a heterogeneidade que pode estar presente neste conjunto de dados (Tabela 6).

Tabela 6 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável altura para a safra 2020/21 em Capinópolis.

	Modelo 1	Modelo 2	Modelo 3
σ_g^2	190,53942	330,17964	208,09909
σ_b^2	92,20846	-	85,88743
σ_r^2	222,50744	-	-
R	-	223,47128	212,09893
h_g^2	0,28757	0,47000	0,31377
r_{gg}	0,53625	0,68557	0,56015
AIC	1987,84299	1994,73307	1990,33875
BIC	1998,75805	2009,28649	2008,53052

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Informatin Criterion* e BIC = *Bayesian Information Criteria*.

Fonte: Elaboração própria.

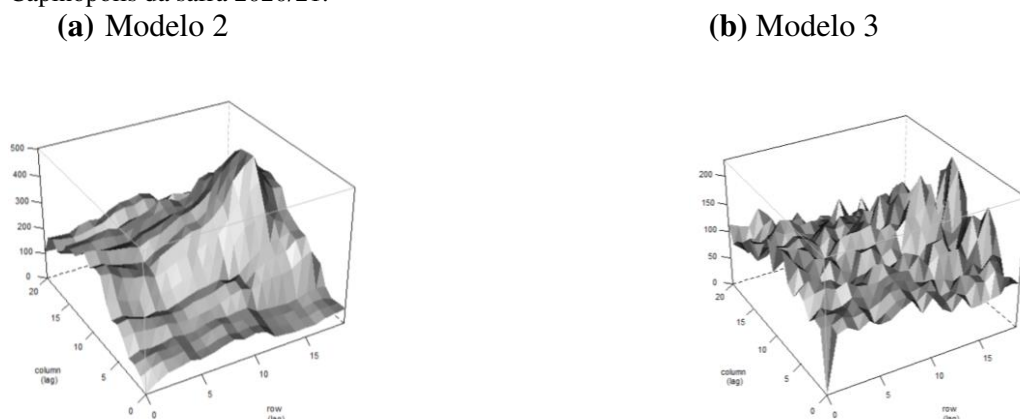
Apesar de não ser necessário considerar a dependência espacial, é interessante observar os variogramas para as duas análises espaciais (Figura 8). No primeiro variograma, em que não foi considerado o efeito de bloco, observa-se maior heterogeneidade na área (ondulação em área total), já quando se considera o bloco no segundo variograma, essa heterogeneidade é reduzida pela blocagem. Assim, visualiza-se que os blocos tenham sido alocado na parte mais heterogênea da área, com variação dentro do mesmo mais homogênea, o que refletiu na análise estatística mais eficiente quanto ao princípio de controle local no modelo considerando a estrutura de DBA.

Tabela 7 – Índice de coincidência (IC) entre os genótipos selecionados para a variável altura para cada modelo estudado e para a safra 2020/21 em Capinópolis.

Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
73%	86%	76%

Fonte: Elaboração própria.

Figura 8 – Figura 7 – Variograma da variável altura para os modelo 2 (a) e modelo 3 (b) do ensaio de Capinópolis da safra 2020/21.



Fonte: Elaboração própria.

Tabela 8 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à altura para Capinópolis safra 2020/21.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC1312S2	94,0617	MDC292S6	115,1885	MDC292S6	96,0339
2	MDC292S6	92,9130	MDC16S5	108,2366	MDC1312S2	94,7173
3	MDC723S1	91,3814	MDC246S3	106,8751	MDC723S1	93,6079
4	MDC152S2	90,0173	MDC723S1	103,2566	MDC152S2	90,6278
5	MDC292S3	87,7951	MDC16S4	102,3835	MDC16S4	89,5486
6	MDC16S4	87,1697	MDC141S6	100,0368	MDC292S3	89,3577
7	MDC25101S1	86,7682	MDC25101S3	98,6418	MDC255S1	88,4368
8	MDC17S2	86,4112	MDC1312S2	97,9966	MDC1312S7	88,2917
9	MDC712S1	86,4112	MDC152S2	97,7334	MDC17S2	88,2380
10	MDC255S1	86,2470	MDC255S1	97,3124	MDC75S6	87,1702
11	MDC92S4	86,2470	MDC1312S7	97,0919	MDC712S1	86,9972
12	MDC75S6	85,9932	MDC18S7	96,7036	MDC25101S1	86,9386
13	MDC84S7	85,9932	MDC17S2	96,5181	MDC261S7	86,3127
14	MDC1312S7	85,9157	MDC75S6	96,1512	MDC84S7	86,3018
15	MDC292S1	85,2967	MDC292S3	95,9368	MDC139S3	85,8304
16	MDC139S3	84,9931	MDC261S7	95,7974	MDC292S1	85,2214
17	MDC261S7	84,4815	MDC723S2	94,2555	MDC18S7	85,1028
18	MDC132S5	84,4617	MDC15S3	92,9412	MDC751S3	85,0775
19	MDC14S1	84,3740	MDC136S4	92,1333	MDC16S5	84,9489
20	MDC152S7	84,1907	MDC751S7	91,3826	MDC92S3	84,9190
21	MDC715S5	84,0705	MDC715S5	91,2888	MDC15S3	84,7299
22	MDC92S3	83,6867	MDC257S4	91,1640	MDC715S5	84,6471

23	MDC152S1	83,5589	MDC723S4	91,0763	MDC18S3	84,6261
24	MDC92S5	83,5589	MDC92S3	90,9049	MDC92S4	84,5858
25	MDC18S3	83,4514	MDC718S6	90,7963	MDC132S5	84,5623
26	MDC751S3	83,2681	MDC25101S1	90,5278	MDC723S4	84,4466
27	MDC751S5	83,0777	MDC710S4	89,9778	MDC152S7	84,2998
28	MDC15S3	83,0178	MDC712S1	89,9286	MDC152S1	84,2511
29	MDC723S4	82,7207	MDC139S3	89,8866	MDC92S5	84,2494
30	MDC716S4	82,6363	MDC751S3	89,8002	MDC25101S3	84,1998
31	MDC18S7	82,5565	MDC84S7	89,4991	MDC710S4	83,9619
32	MDC710S4	82,3455	MDC18S3	89,3036	MDC723S2	83,7323
33	MDC86S3	82,3455	MDC152S7	89,1658	MDC141S6	83,5668
34	MDC71S3	82,1750	MDC92S5	88,7032	MDC14S1	83,4202
35	MDC723S2	82,1551	MDC139S4	88,5411	MDC246S3	83,3140
36	MDC81S5	82,0675	MDC712S2	88,3495	MDC751S7	82,9908
37	MDC110S2	81,6062	MDC152S1	88,0836	MDC751S5	82,6320

Fonte: Elaboração própria.

Quanto ao ciclo, foi encontrada diferença significativa somente entre os modelos 1 e 2 ($p < 0,05$), ou seja, o DBA e Análise Espacial se diferem (Tabela 9). Os valores de AIC e BIC foram menores para para DBA (Tabela 10), o que permite concluir que este foi o modelo que mais se ajusta para estes dados.

Nota-se que, como para altura de planta, para a característica ciclo os valores de σ_r^2 tanto para DBA quanto para análise espacial foram próximos, o que permite inferir que o bloco foi eficiente em capturar possível heterogeneidade existente no ensaio, não sendo necessário considerar correção espacial. Além disso, os valores para σ_b^2 foram muito baixos, como pode ser observado nos variogramas para os modelos de análise espacial (Figura 9). Embora o modelo 2 não considere efeito de bloco e o modelo 3 utilize o efeito, é notável que os dois variogramas não se diferem, ou seja, o controle local por meio da blocagem pode ser dispensada para essas análises.

Tabela 9 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável ciclo para Capinópolis safra 2020/21.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Ciclo	Modelo 1 x Modelo 2	3,3529	0,0335*
	Modelo 1 x Modelo 3	33521,0000	0,08034 ^{ns}
	Modelo 2 x Modelo 3	-0,0008	0,5000 ^{ns}

LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Tabela 10 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável ciclo para a safra 2020/21 em Capinópolis.

	Modelo 1	Modelo 2	Modelo 3
σ_g^2	0,02880	0,02941	0,02880
σ_b^2	0,00448	-	0,00383
σ_r^2	0,00678	-	-
R	-	0,01061	0,00721
h_g^2	0,50810	0,52366	0,51143
r_{gg}	0,71281	0,72364	0,71515
	-	-	-
AIC	686,93395	682,85152	683,83757
	-	-	-
BIC	676,01889	668,29810	665,64580

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Information Criterion* e BIC = *Bayesian Information Criteria*.

Fonte: Elaboração própria.

Tabela 11 – Índice de coincidência (IC) entre os genótipos selecionados para a variável altura para cada modelo estudado e para a safra 2020/21 em Capinópolis.

Ciclo (dias)		
Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
73%	5%	3%

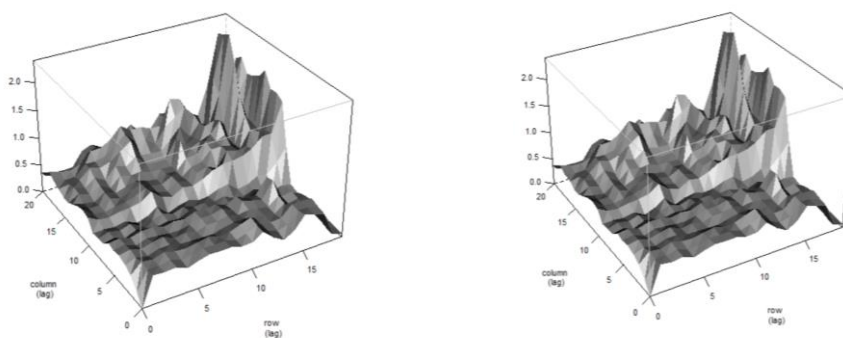
LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Figura 9 – Figura 7 – Variograma da variável ciclo para os modelo 2 (a) e modelo 3 (b) do ensaio de Capinópolis da safra 2020/21.

(a) Modelo 2

(b) Modelo 3



Fonte: Elaboração própria.

Tabela 12 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à ciclo para Capinópolis safra 2020/21.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC18S7	112,2839	MDC18S7	112,2938	MDC152S5	113,4115
2	MDC152S1	112,2557	MDC152S1	112,2448	MDC15S3	112,9721
3	MDC86S3	112,2475	MDC16S4	112,2203	MDC255S1	112,9552
4	MDC17S2	112,2109	MDC17S2	112,2095	MDC266S1	112,9093
5	MDC292S1	112,2007	MDC292S6	112,2038	MDC248S5	112,6175
6	MDC292S6	112,1995	MDC261S7	112,2017	MDC132S3	112,6155
7	MDC92S5	112,1990	MDC266S8	112,2002	MDC291S1	112,6119
8	MDC25101S1	112,1847	MDC751S3	112,1965	MDC292S2	112,6093
9	MDC712S3	112,1833	MDC712S3	112,1923	MDC76S2	112,6036
10	MDC84S7	112,1833	MDC86S3	112,1918	MDC13101S2	112,6029
11	MDC139S3	112,1806	MDC255S1	112,1835	MDC1312S3	112,6026
12	MDC712S1	112,1785	MDC1312S7	112,1812	MDC19S4	112,5742
13	MDC266S8	112,1766	MDC723S4	112,1795	MDC137S2	112,5667
14	MDC751S3	112,1747	MDC751S7	112,1790	MDC136S3	112,5660
15	MDC93S4	112,1747	MDC75S5	112,1741	MDC11S1	112,5649
16	MDC723S4	112,1704	MDC84S7	112,1616	MDC136S2	112,5610
17	MDC14S1	112,1683	MDC17S5	112,1493	MDC86S4	112,5606
18	MDC132S5	112,1604	MDC710S4	112,1493	MDC711S6	112,5605
19	MDC261S7	112,1585	MDC92S5	112,1491	MDC15S4	112,5603
20	MDC723S7	112,1510	MDC712S1	112,1480	MDC82S6	112,5595
21	MDC1312S7	112,1482	MDC723S7	112,1458	MDC19S3	112,5594
22	MDC71S3	112,1423	MDC710S6	112,1417	MDC241S1	112,5557
23	MDC141S5	112,1401	MDC18S1	112,1400	MDC82S2	112,5498
24	MDC92S4	112,1382	MDC710S1	112,1393	MDC85S1	112,5494
25	MDC241S5	112,1380	MDC16S5	112,1386	MDC13101S1	112,5489
26	MDC292S3	112,1380	MDC25101S3	112,1383	MDC248S1	112,5488
27	MDC84S1	112,1380	MDC139S3	112,1375	MDC13S5	112,5481
28	MDC710S4	112,1342	MDC84S1	112,1371	MDC261S2	112,5472
29	MDC16S2	112,1301	MDC712S2	112,1366	MDC75S4	112,5470
30	MDC17S5	112,1267	MDC93S4	112,1336	MDC261S6	112,5450
31	MDC710S1	112,1262	MDC292S1	112,1326	MDC85S2	112,5439
32	MDC16S4	112,1220	MDC25101S1	112,1322	MDC74S3	112,5437
33	MDC255S1	112,1220	MDC25101S2	112,1319	MDC15S1	112,5434
34	MDC710S6	112,1200	MDC132S5	112,1246	MDC141S5	112,5354
35	MDC110S2	112,1197	MDC292S7	112,1240	MDC13S3	112,5286
36	MDC719S1	112,1137	MDC719S3	112,1237	MDC245S4	112,5215
37	MDC18S3	112,1116	MDC710S2	112,1200	MDC717S5	112,5190

Fonte: Elaboração própria.

3.1.2 Safra 2021/22

Para esta safra, para a variável produção, houve diferença quando os modelos 1 e 2 foram comparados e quando os modelos 1 e 3 foram comparados ($p < 0,001$). Para a comparação entre os modelos 2 e 3 não houve diferença significativa ($p > 0,05$) (Tabela 13). Os valores de σ_b^2 encontrados via os modelos 1 e 3 foram baixos, o que sugere baixa heterogeneidade no ensaio. Já os valores das variâncias do resíduo foram mais altos para a análise espacial em comparação ao obtido para o DBA, o que sugere a existência de correlação espacial entre as parcelas (Tabela 14).

Tabela 13 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável produção para Capinópolis safra 2021/22.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Produção	Modelo 1 e Modelo 2	16,8070	0,0000207***
	Modelo 1 e Modelo 3	16,8070	0,0000767***
	Modelo 2 e Modelo 3	-0,1225	0,5000 ^{ns}

Fonte: Elaboração própria.

Estes resultados são concordantes com os valores obtidos para AIC e DBC para este conjunto de dados (Tabela 14). Os menores valores para o modelo 2, que considera dependência espacial sem o efeito de bloco, foram menores, portanto, este modelo foi o que melhor se ajustou para essa característica. Contudo, ao analisar os variogramas apresentados na Figura 10, fica evidenciado que não há diferença entre o modelo que considera correlação espacial (modelo 2) e o modelo que considera correlação espacial e bloco (modelo 3), o que evidencia os valores baixos encontrados para as σ_b^2 .

Tabela 14 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável produção para a safra 2021/22 em Capinópolis.

	Modelo 1	Modelo 2	Modelo 3
σ_g^2	0,0042	26270,7312	26271,0100
σ_b^2	43757,8500	-	0,0054
σ_r^2	41200,7700	-	-
R	-	53611,1756	53611,9500
h_g^2	0,4035	0,3109	0,3109
r_{gg}	0,6352	0,5576	0,5576
AIC	1995,3626	1980,5556	1982,5556
BIC	2004,6254	1992,9059	1997,9936

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Informatin Criterion* e BIC = *Bayesian Information Criteria*.

Tabela 15 - Índice de coincidência (IC) entre os genótipos selecionados para a variável produção para cada modelo estudado e para a safra 2021/22 em Capinópolis.

Produção (g)		
Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
71%	71%	100%

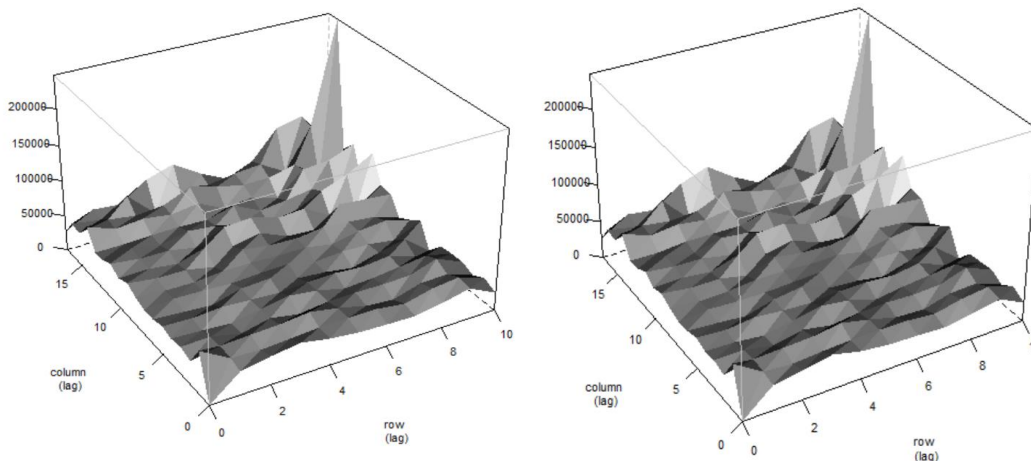
LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Figura 10 – Variograma da variável produção para os modelo 2 (a) e modelo 3 (b) do ensaio de Capinópolis da safra 2021/22.

(a) Modelo 2

(b) Modelo 3



Fonte: Elaboração própria.

Tabela 16 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à produção para Capinópolis safra 2021/22.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC13S3	1215,698076	MDC13S3	1031,56229	MDC13S3	1031,5546
2	MDC246S3	1130,197184	MDC246S3	969,7977443	MDC246S3	969,7878
3	MDC86S1	1043,666579	MDC1312S7	946,6169137	MDC1312S7	946,6105
4	MDC79S3	1016,88335	MDC86S1	910,8072449	MDC86S1	910,8048
5	MDC1312S7	1012,762774	MDC137S3	908,7972226	MDC137S3	908,7901
6	MDC137S3	982,8886694	MDC79S3	875,8876185	MDC79S3	875,8854
7	MDC19S4	969,497139	MDC19S4	869,8661196	MDC19S4	869,8630
8	MDC13S2	961,2563515	MDC19S2	869,0823403	MDC19S2	869,0851
9	MDC19S2	922,1115894	MDC13S2	851,3016234	MDC13S2	851,2976
10	MDC13101S3	908,7195232	MDC13101S3	843,686747	MDC13101S3	843,6826
11	MDC141S5	892,2379056	MDC15S3	837,1743994	MDC15S3	837,1731
12	MDC15S3	873,695554	MDC15S4	834,6647215	MDC15S4	834,6626
13	MDC246S4	854,1236168	MDC136S2	832,5424806	MDC136S2	832,5418
14	MDC292S7	840,7319654	MDC292S7	826,2090511	MDC292S7	826,2059
15	MDC13S7	840,7316279	MDC1312S2	814,3822548	MDC1312S2	814,3834
16	MDC711S5	838,671235	MDC91S4	798,7000753	MDC91S4	798,6913
17	MDC711S6	825,2796486	MDC92S2	783,2690294	MDC92S2	783,2722
18	MDC13101S5	824,2495274	MDC711S6	782,6985094	MDC711S6	782,6956
19	MDC12S1	820,1287197	MDC261S6	782,3818198	MDC261S6	782,3839
20	MDC1312S2	814,9783133	MDC711S5	780,2661695	MDC711S5	780,2631
21	MDC717S5	811,8875825	MDC11S4	777,8414805	MDC11S4	777,8448

Fonte: Elaboração própria.

Para altura de planta foram obtidas diferenças significativas ($p < 0,001$) nas comparações entre DBA e Análise Espacial e entre DBA e Análise Espacial com adição de bloco. Para a comparação entre Análise Espacial e Análise Espacial com adição de bloco não houve diferença significativa ($p > 0,05$) (Tabela 17). Os valores de σ_r^2 foram maiores para os modelos que consideram a dependência espacial, quando comparados com o DBA. No entanto, é possível verificar que os valores para os modelos de análise espacial são próximos, sendo ligeiramente maior para aquele que considera o efeito de bloco (modelo 3). Essa diferença mínima entre os modelos de análise espacial reflete os valores baixos para a σ_b^2 (Tabela 18).

A análise dos critérios AIC e BIC favoreceram selecionar o modelo 2, que considera apenas a análise espacial, como aquele mais eficiente para a característica altura de plantas na maturação (Tabela 18). No entanto, quando analisa-se o variograma, tal diferença entre os modelos 2 e 3 não é perceptível. Isto posto, é importante notar que, apesar de os critérios de comparação evidenciar o melhor ajuste do modelo 2, os valores

de AIC e BIC foram próximos para os três modelos adotados no presente trabalho. Esse fato reforça os resultados obtidos pelo teste de LRT, componentes de variância e as análises dos variogramas (Figura 11).

Tabela 17 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável altura para Capinópolis safra 2021/22.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Altura	Modelo 1 e Modelo 2	11,4530	0,0004***
	Modelo 1 e Modelo 3	11,8280	0,0001***
	Modelo 2 e Modelo 3	0,3742	0,2704 ^{ns}

LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Tabela 18 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável altura para a safra 2021/22 em Capinópolis.

	Modelo 1	Modelo 2	Modelo 3
σ_g^2	6,8664	69,7407	67,0644
σ_b^2	70,0461	-	3,7193
σ_r^2	134,8420	-	-
R	-	145,0635	142,9552
h_g^2	0,2629	0,2979	0,2886
r_{gg}	0,5127	0,5458	0,5372
AIC	1035,0039	1025,5506	1027,1765
BIC	1044,2667	1037,9010	1042,6144

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Informatin Criterion* e BIC = *Bayesian Information Criteria*.

Fonte: Elaboração própria.

Tabela 19 – Índice de coincidência (IC) entre os genótipos selecionados para a variável altura para cada modelo estudado e para a safra 2021/22 em Capinópolis.

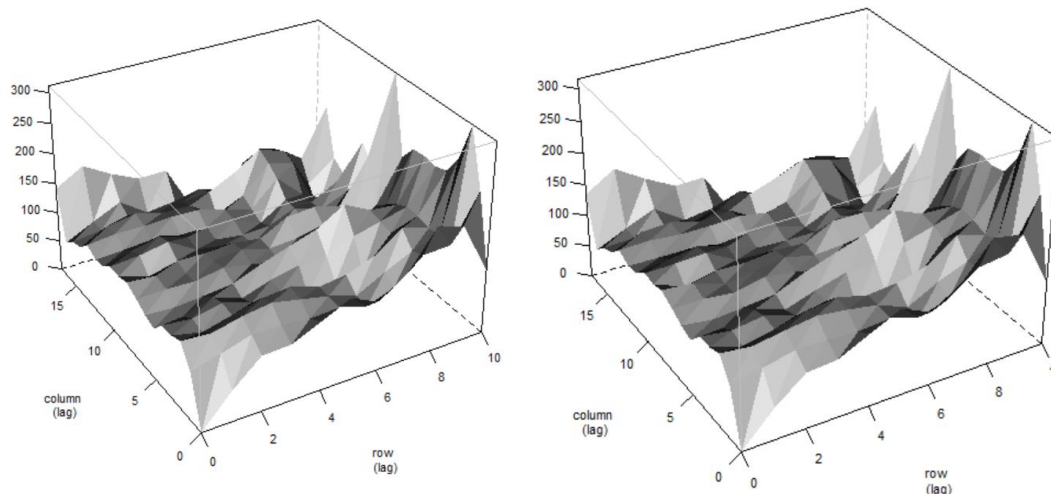
Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
62%	67%	57%

Fonte: Elaboração própria.

Figura 11 –Variograma da variável altura para os modelo 2 (a) e modelo 3 (b) do ensaio de Capinópolis da safra 2021/22.

(a) Modelo 2

(b) Modelo 3



Fonte: Elaboração própria.

Tabela 20 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à altura para Capinópolis safra 2021/22.

Ranking	Modelo 1		Modelo 2		Modelo 3		
	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	1031,5546	MDC245S4	96,9145	MDC245S4	95,7136	MDC245S4	95,4523
2	969,7878	MDC18S7	96,4074	MDC18S7	95,6506	MDC18S7	95,4152
3	946,6105	MDC16S4	94,6962	MDC16S4	95,6052	MDC16S4	95,2343
4	910,8048	MDC13S7	91,9281	MDC13S7	90,8212	MDC13S7	90,5307
5	908,7901	MDC723S6	88,8312	MDC241S1	90,1521	MDC241S1	89,7965
6	875,8854	MDC13S5	87,4791	MDC723S6	89,4606	MDC723S6	89,4865
7	869,8630	MDC13S6	87,4791	MDC110S3	89,0421	MDC85S1	88,7751
8	869,0851	MDC79S2	86,9657	MDC82S6	89,0296	MDC82S6	88,7439
9	851,2976	MDC17S1	86,6475	MDC85S1	88,5503	MDC110S3	88,7232
10	843,6826	MDC74S2	86,6235	MDC74S2	88,0215	MDC74S2	87,7191
11	837,1731	MDC711S6	86,5860	MDC261S4	87,8599	MDC261S4	87,6913
12	834,6626	MDC141S2	86,4355	MDC138S2	87,8560	MDC137S3	87,5839
13	832,5418	MDC85S1	86,4118	MDC137S3	87,8396	MDC138S2	87,5105
14	826,2059	MDC137S3	86,4083	MDC76S4	87,5586	MDC79S2	87,2685
15	814,3834	MDC138S2	86,2813	MDC79S2	87,5283	MDC76S4	87,0287
16	798,6913	MDC74S1	86,2644	MDC138S1	86,8170	MDC138S1	86,8165
17	783,2722	MDC76S4	86,2437	MDC76S2	86,6995	MDC76S2	86,3532
18	782,6956	MDC15S5	86,2362	MDC17S4	86,4747	MDC255S1	86,3430
19	782,3839	MDC261S4	86,0661	MDC13S5	86,2243	MDC241S7	86,1737
20	780,2631	MDC292S3	85,9222	MDC86S1	86,1022	MDC17S4	86,1607
21	777,8448	MDC241S7	85,8984	MDC13S3	86,0890	MDC15S5	85,9049

Fonte: Elaboração própria.

Para a característica ciclo não houve diferença significativa para todos os modelos comparados ($p > 0,05$) (Tabela 21). Nota-se baixo valor para σ_b^2 do modelo 1 e σ_b^2 igual a zero para o modelo 3, o que sugere que os blocos foram eficientemente alocados. Apesar de diferirem pouco, os valores de AIC e BIC foram menores para DBA, fazendo com que este seja o modelo mais ajustado para os dados (Tabela 22). Comparando-se o variograma dos modelos 2 e 3 também pode-se notar a ausência de variação de bloco, já que os variogramas foram similares (Figura 12).

Tabela 21 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável ciclo para Capinópolis safra 2021/22.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Ciclo	Modelo 1 e Modelo 2	0,3910	0,2659 ^{ns}
	Modelo 1 e Modelo 3	0,3910	0,4715 ^{ns}
	Modelo 2 e Modelo 3	-0,0718	0,5000 ^{ns}

LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Tabela 22 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável ciclo para a safra 2021/22 em Capinópolis.

	Ciclo(dias)		
	Modelo 1	Modelo 2	Modelo 3
σ_g^2	0,0856	8,5849	8,5868
σ_b^2	9,0210	-	0,0000
σ_r^2	36,2714	-	-
R	-	36,8364	36,7948
h_g^2	0,1559	0,1501	0,1501
r_{gg}	0,3949	0,3874	0,3874
AIC	794,2242	795,8332	797,8332
BIC	803,4870	808,1835	813,2711

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Information Criterion* e BIC = *Bayesian Information Criteria*.

Fonte: Elaboração própria.

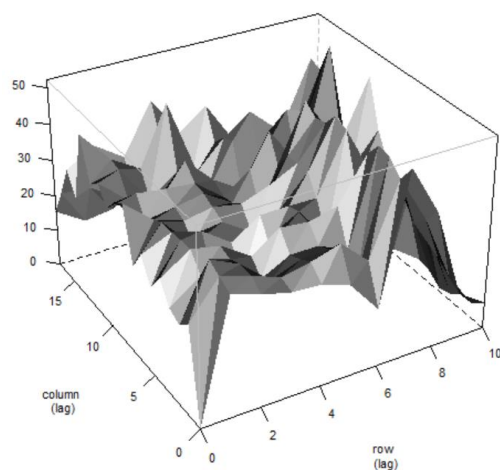
Tabela 23 – Índice de coincidência (IC) entre os genótipos selecionados para a variável ciclo para cada modelo estudado e para a safra 2021/22 em Capinópolis.

Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
52%	52%	100%

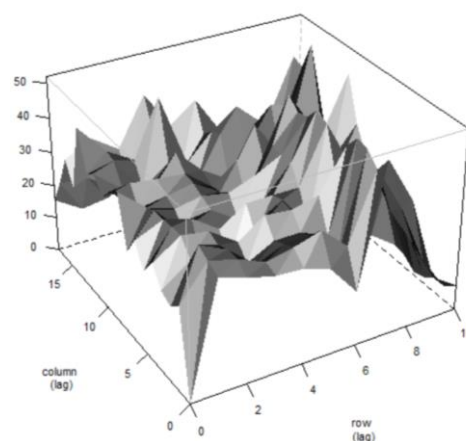
Fonte: Elaboração própria.

Figura 12 – Variograma da variável ciclo para os modelo 2 (a) e modelo 3 (b) do ensaio de Capinópolis da safra 2021/22.

(a) Modelo 2



(b) Modelo 3



Fonte: Elaboração própria.

Tabela 24 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à altura para Capinópolis safra 2021/22.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC11S1	131,4309	MDC152S5	132,1963	MDC152S5	132,2115
2	MDC132S3	131,4309	MDC15S3	131,7629	MDC15S3	131,7721
3	MDC136S3	131,4309	MDC255S1	131,7463	MDC255S1	131,7552
4	MDC15S3	131,4309	MDC266S1	131,7009	MDC266S1	131,7093
5	MDC241S1	131,4309	MDC248S5	131,4122	MDC248S5	131,4175
6	MDC246S3	131,4309	MDC132S3	131,4109	MDC132S3	131,4155
7	MDC261S2	131,4309	MDC291S1	131,4095	MDC291S1	131,4119
8	MDC292S1	131,4309	MDC292S2	131,4084	MDC292S2	131,4093
9	MDC76S2	131,4309	MDC76S2	131,3991	MDC76S2	131,4036
10	MDC85S3	131,4309	MDC1312S3	131,3987	MDC13101S2	131,4029
11	MDC12S1	131,4279	MDC13101S2	131,3982	MDC1312S3	131,4026
12	MDC13101S2	131,4279	MDC19S4	131,3695	MDC19S4	131,3742
13	MDC141S2	131,4279	MDC11S1	131,3633	MDC137S2	131,3667
14	MDC152S5	131,4279	MDC137S2	131,3620	MDC136S3	131,3660
15	MDC19S4	131,4279	MDC136S3	131,3617	MDC11S1	131,3649
16	MDC292S3	131,4279	MDC136S2	131,3566	MDC136S2	131,3610
17	MDC86S4	131,4279	MDC711S6	131,3563	MDC86S4	131,3606
18	MDC136S2	131,4206	MDC86S4	131,3561	MDC711S6	131,3605
19	MDC18S2	131,4206	MDC15S4	131,3559	MDC15S4	131,3603
20	MDC19S3	131,4206	MDC82S6	131,3551	MDC82S6	131,3595
21	MDC241S7	131,4206	MDC19S3	131,3551	MDC19S3	131,3594

Fonte: Elaboração própria.

3.2 Ensaio de Madre de Deus

Para os ensaios conduzidos em Madre de Deus de Minas, apenas a característica produção de grãos foi utilizada para as análises para ambas as safras.

3.2.1 Safra 2019/20

Para esta safra, o teste LRT apontou diferença significativa entre os modelos 1 e 2 ($P < 0,001$) e entre os modelos 1 e 3 ($p < 0,01$). Não foi encontrada diferença significativa para a comparação entre modelos 2 e 3 ($p > 0,05$) (Tabela 25). A variância de bloco para o modelo de DBA apresentou maior valor em detrimento aos apresentados pelos modelos baseados na análise espacial (Tabela 26). Estes resultados sugerem que os blocos foram alocados de forma ineficiente na área, fazendo com que a blocagem experimental não seja eficiente em neutralizar os efeitos da heterogeneidade da área. Deste modo, faz-se necessário considerar o uso de um modelo que considere a correlação espacial entre parcelas.

Tabela 25 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável produção para Madre de Deus safra 2019/20.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Produção	Modelo 1 x Modelo 2	6,4572	0.0055 **
	Modelo 1 x Modelo 3	6,4572	0.0154 *
	Modelo 2 x Modelo 3	-0,0458	0.5000ns

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Produção	Modelo 1 x Modelo 2	10,1710	0.0007 ***
	Modelo 1 x Modelo 3	10,1710	0.0023 **
	Modelo 2 x Modelo 3	0,0151	0.4984 ^{ns}

Fonte: Elaboração própria.

Os valores da variância do resíduo para os modelos de análise espacial foram maiores quando comparados com o apresentado para o modelo que utiliza a estrutura do DBA (Tabela 25). De acordo com o exposto, os valores de AIC e BIC indicaram que o modelo 2 como o de melhor ajuste aos dados para produção de grãos (Tabela 26). Nota-se também que a análise espacial conseguiu corrigir grandemente a variância entre blocos, captando essa heterogeneidade para o resíduo, o que é possível verificar pela similaridade

dos variogramas (Figura 13) e pela pouca diferença entre AICe BIC em ambos os modelos.

Tabela 26 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável produção para a safra 2019/20 em Madre de Deus.

	Produção (g)		
	Modelo 1	Modelo 2	Modelo 3
σ_g^2	634,2049	30373,3900	30355,4900
σ_b^2	34387,8255	-	0,0016
σ_r^2	36159,8556	-	-
r	-	39391,1500	39381,9900
h ²	0,4035	0,3109	0,3109
r _{gg}	0,6352	0,5576	0,5576
AIC	1995,3626	1980,5556	1982,5556
BIC	2004,6254	1992,9059	1997,9936

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h²_g = herdabilidade de Cullis; AIC = Akaike Information Criterion e BIC = Bayesian Information Criteria.

Fonte: Elaboração própria.

Tabela 27 – Índice de coincidência (IC) entre os genótipos selecionados para a variável produção para cada modelo estudado e para a safra 2019/20 em Capinópolis.

Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
91%	91%	100%

Fonte: Elaboração própria.

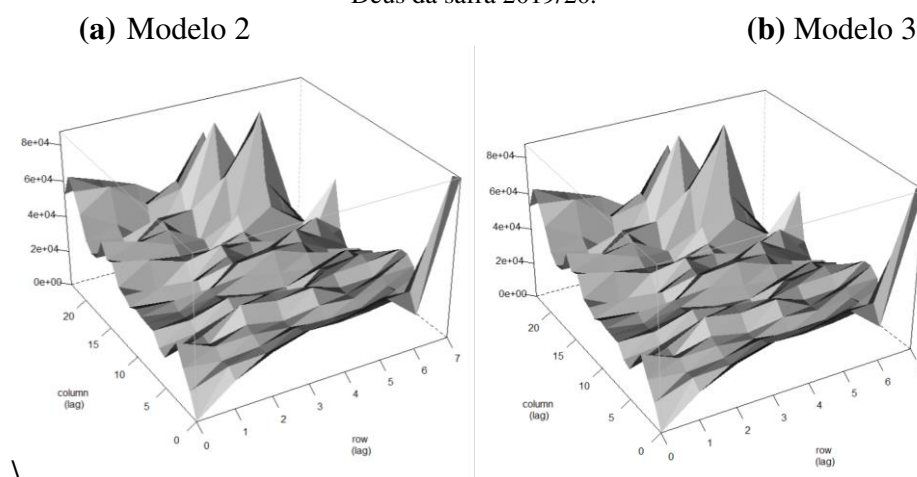
Tabela 28 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à altura para Madre de Deus safra 2019/20.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC137S3	436,6359	MDC137S3	445,7152	MDC137S3	446,1063
2	MDC18S1	282,9433	MDC18S1	280,3138	MDC18S1	280,6023
3	MDC17S4	245,9105	MDC17S4	221,8693	MDC17S4	221,8054
4	MDC717S5	161,0478	MDC717S5	216,1831	MDC717S5	216,7450
5	MDC110S3	146,0874	MDC110S3	165,4753	MDC110S3	165,1614
6	MDC292S2	142,1029	MDC17S2	152,8325	MDC17S2	153,1271
7	MDC17S2	132,6590	MDC18S4	150,6048	MDC18S4	151,1238
8	MDC79S3	120,9560	MDC292S2	143,1035	MDC292S2	143,0578
9	MDC93S4	116,0604	MDC13S6	130,1796	MDC13S6	130,5098
10	MDC1312S4	109,2272	MDC19S1	129,4473	MDC19S1	128,9734
11	MDC18S4	108,9813	MDC13101S3	120,7102	MDC13101S3	120,2370
12	MDC110S1	107,1753	MDC81S2	107,0394	MDC81S2	107,3638
13	MDC81S2	101,7275	MDC110S1	101,2816	MDC110S1	101,3284
14	MDC72S2	95,7132	MDC93S4	97,3123	MDC93S4	97,1952
15	MDC13S6	94,6366	MDC79S3	96,7827	MDC79S3	96,8810
16	MDC241S6	87,9539	MDC1312S4	89,4828	MDC1312S4	89,4477

17	MDC266S5	87,4671	MDC13S3	86,5877	MDC13S3	86,6166
18	MDC19S1	87,1051	MDC723S6	80,2051	MDC723S6	80,3200
19	MDC72S7	84,7324	MDC15S5	76,5382	MDC15S5	76,6799
20	MDC15S5	81,1223	MDC241S6	75,8898	MDC241S6	75,9049
21	MDC13101S3	79,3967	MDC72S7	70,3689	MDC72S7	70,2346
22	MDC261S4	79,2700	MDC85S2	62,3809	MDC85S2	62,1629
23	MDC13S3	75,0355	MDC72S2	61,9938	MDC72S2	61,8494

Fonte: Elaboração própria.

Figura 13 – Variograma da variável ciclo para o modelo 2 (a) e modelo 3 (b) do ensaio de Madre de Deus da safra 2019/20.



Fonte: Elaboração própria.

3.2.2 Safra 2020/21

Para esta safra houve diferença significativa entre os modelos 1 e 2 ($p < 0,001$) e entre os modelos 1 e 3 ($p < 0,01$) (Tabela 29). Não houve diferença significativa quando os modelos 2 e 3 foram comparados ($p > 0,05$). Os resultados para AIC e BIC apontam o modelo 2 como o de melhor ajuste (Tabela 30). Por meio do variograma (Figura 13) também foi possível verificar a similaridade entre o ajuste dos modelos que consideram a correlação espacial.

Tabela 29 – Teste de razão de verossimilhança (LRT) dos modelos testados para a variável produção para Madre de Deus safra 2019/20.

Variável	Modelo Reduzido x Modelo Completo	LRT	Pr (χ^2)
Produção	Modelo 1 x Modelo 2	10,1710	0.0007 ***
	Modelo 1 x Modelo 3	10,1710	0.0023 **
	Modelo 2 x Modelo 3	0,0151	0.4984 ^{ns}

LRT = *Likelihood Ratio Test*; Pr = Probabilidade pelo teste de qui-quadrado; ns = não significativo; *significativo a 5% de probabilidade; **significativo a 1% de probabilidade e ***significativo a 0,1% de probabilidade.

Fonte: Elaboração própria.

Tabela 30 – Componentes de variância, parâmetros genéticos e índices de comparação para os modelos estudados para a variável produção para a safra 2021/22 em Madre de Deus.

	Modelo 1	Modelo 2	Modelo 3
σ_g^2	0,0019	12200,8338	12200,0300
σ_b^2	12656,1000	-	0,0019
σ_r^2	18395,3200	-	-
R	-	18710,0455	18709,7600
h_g^2	0,3302	0,3530	0,3530
r_{gg}	0,5746	0,5942	0,5941
AIC	1896,8687	1888,6975	1890,6974
BIC	1906,2227	1901,1694	1906,2874

σ_g^2 = variância genotípica; σ_b^2 = variância de bloco; σ_r^2 = variância do resíduo; R = resíduo considerando correlação espacial; r_{gg} = acurácia; h_g^2 = herdabilidade de Cullis; AIC = *Akaike Informatin Criterion* e BIC = *Bayesian Information Criteria*.

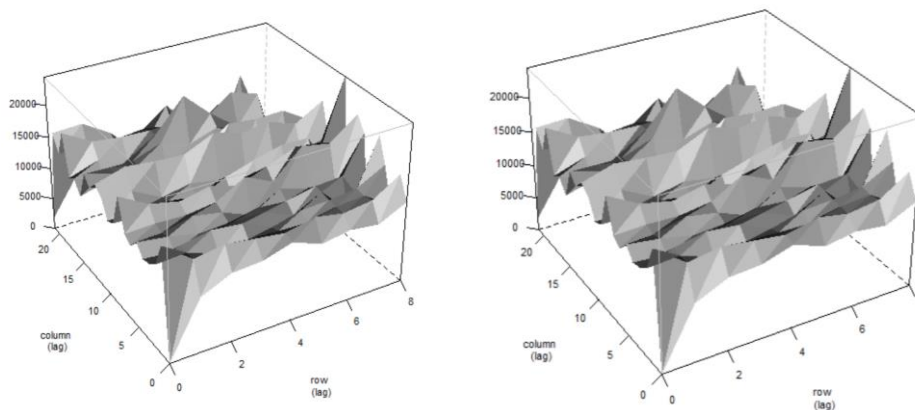
Fonte: Elaboração própria.

Tabela 31 – Índice de coincidência (IC) entre os genótipos selecionados para a variável produção para cada modelo estudado e para a safra 2019/20 em Capinópolis.

Modelo 1 e Modelo 2	Modelo 1 e Modelo 3	Modelo 2 e Modelo 3
86%	86%	100%

Fonte: Elaboração própria.

Figura 14 – Variograma das variáveis produção para os modelos 2 e 3, respectivamente, dos ensaios de Madre de Deus das safras 2021/22.



Fonte: Elaboração própria.

Tabela 32 – Ranking e valores genotípicos dos genótipos selecionados (taxa de seleção = 15%) em cada modelo estudado quanto à altura para Madre de Deus safra 2021/22.

Ranking	Modelo 1		Modelo 2		Modelo 3	
	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico	Genótipos	Valor Genotípico
1	MDC86S1	181,7707	MDC86S1	183,3442	MDC86S1	183,3587
2	MDC13101S5	153,2391	MDC13101S5	180,3170	MDC13101S5	180,3314
3	MDC81S6	147,5320	MDC81S6	147,1148	MDC81S6	147,1251
4	MDC92S4	136,9344	MDC92S4	131,0882	MDC92S4	131,0978
5	MDC14S3	128,7831	MDC241S4	118,1869	MDC241S4	118,1948
6	MDC152S5	120,6311	MDC15S1	110,6352	MDC15S1	110,6453
7	MDC79S3	119,8154	MDC15S3	110,4740	MDC15S3	110,4832
8	MDC246S4	114,1089	MDC15S5	109,4849	MDC15S5	109,4927
9	MDC15S5	108,4025	MDC246S4	104,6468	MDC246S4	104,6513
10	MDC15S1	105,1418	MDC137S3	96,2610	MDC137S3	96,2706
11	MDC292S7	102,6961	MDC292S7	95,7280	MDC292S7	95,7381
12	MDC91S1	96,1745	MDC91S1	94,0757	MDC91S1	94,0872
13	MDC71S8	95,3598	MDC11S1	92,7285	MDC11S1	92,7398
14	MDC12S5	93,7295	MDC110S3	90,1604	MDC110S3	90,1675
15	MDC132S3	91,2832	MDC14S3	87,5865	MDC14S3	87,5841
16	MDC15S3	90,4680	MDC79S3	86,2963	MDC79S3	86,2989
17	MDC241S4	89,6531	MDC71S8	85,7558	MDC71S8	85,7613
18	MDC19S5	87,2072	MDC132S3	85,0611	MDC132S3	85,0650
19	MDC110S3	83,1312	MDC152S5	82,6345	MDC152S5	82,6398
20	MDC1312S7	79,0558	MDC138S5	81,8147	MDC138S5	81,8232
21	MDC241S7	76,6102	MDC241S7	76,8250	MDC241S7	76,8295

Fonte: Elaboração própria.

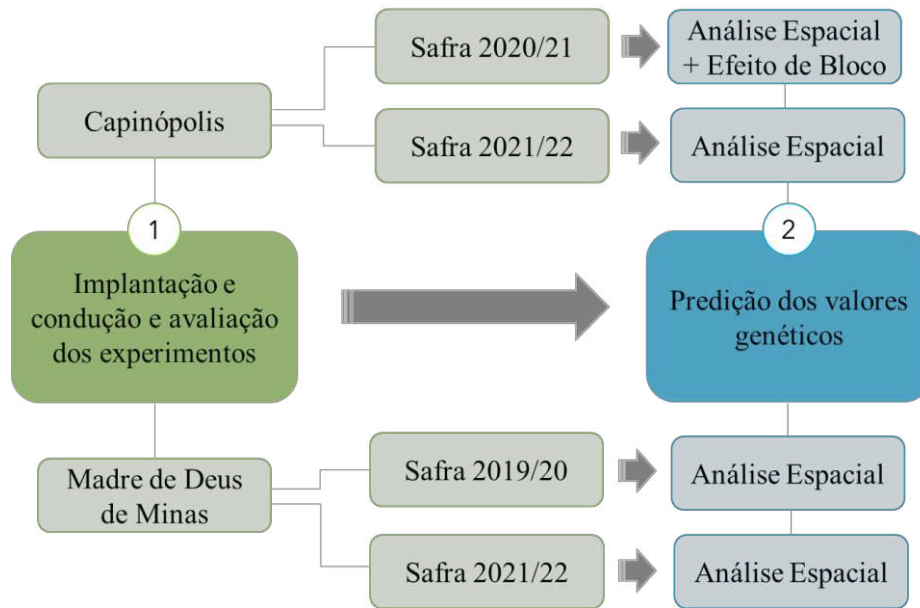
Costa *et al.* (2001), em um estudo que compara o delineamento em blocos casualizados e a abordagem da análise espacial para espécies florestais, concluiu que a análise que inclui a dependência espacial aumenta sobremaneira a eficiência das análises. Qiao *et al.* (2000) identificou em estudo para obtenção de valores genéticos em ensaios de trigo um aumento na eficiência de 14% na predição dos valores para modelos de análise espacial.

4. CONCLUSÕES

Tendo em vista que uma estimativa precisa do mérito genético dos genótipos é essencial para o sucesso de um programa de melhoramento e que o uso de modelos mais ajustados para a realização das análises é imprescindível, para este estudo, conclui-se que os modelos que melhor se ajustaram foram os que consideram a correlação espacial entre as parcelas. Para os ambientes Capinópolis safra 2021/22, Madre de Deus Safra 2019/20

e safra 2021/22 foi preferencial a utilização da análise espacial sem a adição de blocos. Para o ambiente Capinópolis safra 2020/21 é preferível o modelo que abrange a análise espacial com a adição de blocos (Figura 15).

Figura 15 - Fluxograma das conclusões apresentadas para os ensaios realizados em Capinópolis (MG) e Madre de Deus (MG).



REFERÊNCIAS

- AKAIKE, H. **Stochastic theory of minimal realization**, this issue, pp. 667–674.
- AZEVEDO, Cleber Vinicius Giaretta *et al.* Genetic parameters of soybean populations obtained from crosses between grain and food genotypes. **Acta Scientiarum. Agronomy**, v. 43, 2020.
- BASAVARAJA, Talavar *et al.* Assessment of phenotypic diversity and multi-locational screening against bean common mosaic virus (BCMV) disease resistance in dry bean (*Phaseolus vulgaris* L.) germplasm. **Plant Genetic Resources**, p. 1-8, 2022.
- BERNARDELI, Arthur *et al.* Modeling spatial trends and enhancing genetic selection: An approach to soybean seed composition breeding. **Crop Science**, v. 61, n. 2, p. 976-988, 2021
- BRENNAN, J. P. (1988). An economic investigation of wheat breeding programmes. **Agricultural Economics Bulletin** 35, Department of Agricultural Economics and Business Management, **University of New England, Armidale, NSW, Australia**.
- CARNEIRO, Monalisa Sampaio *et al.* RB985476-a sugarcane cultivar with high agro-industrial yield and disease resistance. **Crop Breeding and Applied Biotechnology**, v. 20, 2020.
- CULLIS, B.R.; GLEESON, A.C. Spatial analysis of field experiments – an extension at two dimensions. **Biometrics**, v. 47, p. 1449 – 1460, 1991.
- CULLIS, B.R.; SMITH, A.B.; COOMBES, N.E. 2006. On the design of early generation variety trials with correlated data. **J. Agric. Biol. Environ. Stat.** 11(4): 381–393. doi: 10.1198/108571106X154443.
- DUARTE, João Batista; VENCOVSKY, Roland; DIAS, Carlos Tadeu dos Santos. Estimadores de componentes de variância em delineamento de blocos aumentados com tratamentos novos de uma ou mais populações. **Pesquisa Agropecuária Brasileira**, v. 36, p. 1155-1167, 2001.
- DUARTE, João Batista; VENCOVSKY, Roland; DIAS, Carlos Tadeu dos Santos. Estimadores de componentes de variância em delineamento de blocos aumentados com tratamentos novos de uma ou mais populações. **Pesquisa Agropecuária Brasileira**, v. 36, p. 1155-1167, 2001.
- EISENBERG, B.E.; GAUCH, H.G.; ZOBEL, R.W.; KILIAN, W. Spatial analysis of field experiments: fertilizer experiments with wheat (*Triticum aestivum*) and tea (*Camellia sinensis*). In: KANG, M.S.; GAUCH, H.G. (Ed.) **Genotype by environment interaction**. Boca Raton: **CRC Press**, 1996. Cap. 14, p. 373-404.
- ELIAS, A. A. *et al.* Improving Genomic Prediction in Cassava Field Experiments Using Spatial Analysis. **G3: Genes, Genomes, Genetics**. V. 8, 2018.
- FEDERER, W.T. Augmented (or hoonuiaku) designs. **Hawaiian Planter's Records**, v. 55, p. 191 – 208, 1956.
- FEDERER, W.T. Augmented designs with one-way elimination of heterogeneity. **Biometrics**, v. 17, p. 447 – 473, 1961a.

FEDERER, W.T. Augmented designs with two-, three- and higher-way elimination of heterogeneity. **Biometrics**, v. 17, p. 166, 1961b.

FEDERER, W.T. **Augmented designs**. **Biometrics**, v. 14, p. 134 (Abstract 469), Mar 1958.

FEDERER, W.T.; RAGHAVARAO, D. On augmented designs. **Biometrics**, v. 31, p. 29 – 35, 1975.

GAHLAUT, Vijay *et al.* Dataset of nine agronomic traits in bread wheat phenotyped under irrigated and rain-fed environments. **Data in Brief**, v. 41, p. 107933, 2022.

GILMOUR, A.R.; GOGEL, B.J.; CULLIS, B.R., WELHAM, S. J.; THOMPSON, R. 2015. ASReml user guide release 4.1 structural specification. Hemel Hempstead VSN Int. Ltd. R Core Team. 2019. R: A Language and Environment for Statistical Computing (Version 3.5.2, R Foundation for Statistical Computing, Vienna, Austria, 2018).

GOGEL, Bevrley *et al.* Empirical comparison of time series models and tensor product penalised splines for modeling spatial dependence in plant breeding field trials. **Frontiers in Plant Science**, v. 13, 2023.

HENDERSON, C.R. 1975. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics** 31(2): 423. doi: 10.2307/2529430.

ISLAM, Md S. *et al.* Experimental evaluation of genomic selection prediction for rust resistance in sugarcane. **The Plant Genome**, v. 14, n. 3, p. e20148, 2021.

JIMÉNEZ-GALINDO, José Cruz *et al.* Mapping of resistance to corn borers in a MAGIC population of maize. **BMC plant biology**, v. 19, n. 1, p. 1-17, 2019.

PATTERSON, H.D.; THOMPSON, R. 1971. Recovery of inter-block information when block sizes are unequal. **Biometrika** 58(3): 545–554.

PUTRI, P. H. *et al.* Genetic parameters of soybean [*Glycine Max (L.)*] Genotypes tolerant to salinity. Legume **Research-An International Journal**, v. 45, n. 6, p. 695-699, 2022.

RESENDE, M. D. V.; STURION, J. A. Análise estatística espacial de experimentos via modelos mistos individuais com erros modelados por processos ARIMA em duas dimensões. **Revista de matemática e estatística**, v. 21, p. 7-33, 2003.

RESENDE, M.D.V.; SILVA, F.F.; AZEVEDO, C.F. 2014. **Estatística matemática, biométrica e computacional**: Modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTLGWAS, estatística espacial e temporal, competição, sobrevivência. Suprema, Visconde do Rio Branco.

RODRÍGUEZ-ÁLVAREZ, María Xosé *et al.* Correcting for spatial heterogeneity in plant breeding experiments with P-splines. **Elsevier**, v. 23, p.52-71, 2018.

SABA, Iram *et al.* Using augmented design for evaluation of common bean (*Phaseolus vulgaris L.*) germplasm. **International Journal of Current Microbiology and Applied Science**, v. 6, n. 7, p. 246-254, 2017.

SALUCI, Júlio Cesar Gradice *et al.* Potential of Popcorn S4 lines for resistance to southern corn leaf blight. **Functional Plant Breeding Journal**, v. 2, n. 2, 2020.

SANGHERA, G. S.; JAMWAL, N. S. **Evaluation of elite sugarcane clones for cane yield and component traits using augmented design.** 2019.

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

SILVA, Michele Jorge *et al.* Experimental precision of spatial analysis methods to evaluate the productivity of common bean families. **Revista de Ciências Agrárias**, v.63, 2020.

WOYANN, Leomar Guilherme *et al.* Selection of high-yielding, adapted and stable wheat lines in preliminary trials. **Crop breeding and applied biotechnology**, v. 19, p. 412-419, 2019.

CAPÍTULO 2

**CONSTRUÇÃO DE IDEÓTIPO EM DIFERENTES CENÁRIOS COM USO DO
MGIDI PARA SELEÇÃO DE LINHAGENS SUPERIORES DE SOJA**

RESUMO

RODRIGUES, Fernanda Cupertino, D.Sc., Universidade Federal de Viçosa, junho de 2024. **Seleção de genótipos superiores de soja via MGIDI**. Orientador: Felipe Lopes da Silva. Coorientadores: Luiz Alexandre Peternelli e Maicon Nardino.

Atualmente, o Brasil é o maior produtor mundial de soja. Em virtude da importância dessa cultura para o Brasil e para o mundo, programas de melhoramento genético procuram selecionar materiais superiores em diversas características para atender ao mercado sojícola. Dado o exposto, o objetivo deste trabalho foi selecionar genótipos superiores para a região de Capinópolis via modelos mistos (REML/BLUP) e índice MGIDI (*Multi-trait Genotype-Ideotype Distance Index*). Para tal, ensaios de linhagens de soja do Programa Soja da Universidade Federal de Viçosa foram implantados em Capinópolis – Minas Gerais nas safras 2020/21 e 2021/22. As variáveis analisadas foram: produção (gramas), altura (cm) de planta e ciclo (dias), para ambas as safras. De posse dos valores genéticos obtidos por meio dos modelos mistos, o MGIDI foi aplicado em diferentes cenários para verificação da melhor estratégia de definição do ideótipo. O ideótipo foi desenhado em seis diferentes cenários: (1) considerando apenas o sentido; (2) considerando o sentido + peso; (3) considerando valores da média da melhor testemunha; (4) considerando valores da média da melhor testemunha + peso; (5) considerando valor de uma cultivar alvo para a região e (6) considerando valor de uma cultivar alvo para a região + peso. Os cenários foram comparados por meio do ganho de seleção (GS), do ganho de seleção ajustado (GSA) e do índice de coincidência (IC). As análises demonstraram que o cenário 1 é o que consegue maior ganho genético total quando consideradas todas as características simultaneamente, no entanto, o cenário 5 é o que consegue maiores ganhos de seleção para produção. Tendo em vista que não foram obtidos ganhos expressivos para ciclo, que a altura de planta da população selecionada já atende ao ideótipo requerido e que o incremento da produção é o maior objetivo de um programa de melhoramento, para este conjunto de dados a melhor construção do ideótipo foi a utilizada no cenário 5. Portanto, os genótipos selecionados foram: MDC1312S7, MDC139S3, MDC25101S2, MDC110S2, MDC292S7, MDC137S3, MDC141S5, MDC241S5, MDC15S5, MDC246S3, MDC13S2, MDC241S6, MDC82S2, MDC248S1, MDC141S6, MDC141S2, MDC81S5, MDC91S4 e MDC255S1.

Palavra(s) chave: *Glycine max* (L.) Merr.; Índice multi característico; MGIDI; Análise espacial; Modelos mistos; Delineamento aumentado

ABSTRACT

RODRIGUES, Fernanda Cupertino, D.Sc., Universidade Federal of Viçosa, June, 2024. **Selection of superior soybean genotypes via MGIDI.** Advisor: Felipe Lopes da Silva. Co-advisors: Luiz Alexandre Peternelli e Maicon Nardino.

Currently, Brazil is the world's largest soybean producer. Due to the importance of this crop for Brazil and the world, breeding programs seek to select superior materials in several characteristics to serve the soybean market. Given the above, the objective of this work was to select superior genotypes for the Capinópolis region via mixed models (REML/BLUP) and MGIDI index (Multi-trait Genotype–Ideotype Distance Index). To this end, trials of soybean lines from the Soybean Program of the Federal University of Viçosa were implemented in Capinópolis – Minas Gerais in the 2020/21 and 2021/22 harvests. The variables checked were: production (grams), plant height (cm) and cycle (days), for both harvests. With the genetic values obtained through the mixed models, the MGIDI was applied in different scenarios to verify the best strategy for defining the ideotype. The ideotype was designed in six different scenarios: (1) considering only the meaning; (2) considering meaning + weight; (3) consideration of average values of the best witness; (4) considering average values of the best witness + weight; (5) considering the value of a target cultivar for the region and (6) considering the value of a target cultivar for the region + weight. The scenarios were compared using the selection gain (GS), the adjusted selection gain (GSA) and the coincidence index (CI). The analyzes demonstrated that the scenario is the one that achieves the greatest total genetic gain when considering all characteristics simultaneously, however, the scenario is the one that achieves the greatest gains in selection for productivity. Considering that no significant gains were obtained for the cycle, that the plant height of the selected population already meets the required ideotype and that increasing production is the main objective of breeding program, for this set of data, the best ideotype construction was used in scenario five. Therefore, the selected genotypes were: MDC1312S7, MDC139S3, MDC25101S2, MDC110S2, MDC292S7, MDC137S3, MDC141S5, MDC241S5, MDC15S5, MDC246S3, MDC13S2, MDC241S6, MDC82S2, MDC248S1, 14 1S6, MDC141S2, MDC81S5, MDC91S4 and MDC255S1.

Keyword(s): *Glycine max* (L.) Merr.; Multi-characteristic index; MGIDI; Spatial analysis; Mixed models

1. INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) é uma espécie autógama originária da Ásia, mais precisamente da região da Manchúria, na Índia. Atualmente, o Brasil é o maior produtor do mundo, neste ano, na safra 23/2024, o Brasil produziu mais de 147 milhões de toneladas e atingiu uma produção de 3.205 kg/ha (CONAB, 2024).

Dada a importância da cultura da soja no cenário nacional e mundial, e para atender às exigências dos produtores, os melhoristas estão sempre em busca de cultivares superiores para serem lançadas nas cinco macrorregiões sojícolas do Brasil. Para tanto, existem algumas ferramentas estatísticas que aumentam a chance de êxito no processo de seleção de genótipos superiores, dentre elas estão os índices de seleção (CRUZ *et al.*, 2012). O incremento da produção é o principal objetivo de um programa de melhoramento, no entanto, é necessário que o material genético selecionado reúna, simultaneamente, uma série de atributos favoráveis que lhe confira rendimento comparativamente mais elevado que satisfaça as exigências do mercado. Alguns destes índices permitem que a seleção seja realizada com base em várias características ao mesmo tempo. Tais índices criam valores numéricos, que funcionam como um caráter adicional à análise, resultante da combinação de determinados caracteres escolhidos pelo melhorista, sobre os quais deseja exercer a seleção simultânea (GRANATE *et al.*, 2002).

Dentre os índices disponíveis na literatura está o MGIDI (*Multi-trait Genotype–Ideotype Distance Index*), proposto por Olivoto & Nardino (2021). Quando comparados com outros índices clássicos, como o Índice Smith-Hazel (SMITH, 1936; HAZEL, 1943), o MGIDI, possui como características vantajosas: i) capacidade de contornar problemas de multicolineariedade; ii) não necessita da atribuição de pesos econômicos; iii) a correlação genética é considerada, permitindo preservar as relações entre as características e fornecendo um significado biológico aos ideótipos (ROCHA *et al.*, 2017; ROCHA *et al.*, 2019; OLIVOTO; NARDINO, 2021; FERREIRA, 2020; NASCIMENTO, 2020; VOLPATO *et al.*, 2021). As análises estatísticas do índice MGIDI são realizadas em quatro principais etapas: i) na primeira, o planejamento de um ideótipo é realizado; ii) posteriormente, a análise redimensiona os valores das características de 0-100; iii) nesta etapa, uma análise exploratória de fatores é utilizada para explicar a estrutura da correlação e a dimensionalidade da redução dos dados; e iv) por último, são calculadas as distâncias entre genótipos e ideótipos (OLIVOTO; NARDINO, 2021).

Desde que o MGIDI foi proposto, programas de melhoramento de diferentes culturas passaram a utilizá-lo como estratégia de seleção de genótipos superiores com base em várias características (BENAKANNAHALLI *et al.*, 2021; POUR-ABOUGHADAREH *et al.*, 2021; POUR-ABOUGHADAREH; POCZAI, 2021; OLIVOTO *et al.*, 2021; DALAROSA, 2021; SOARES, 2022; MOREIRA, 2022; TOSOTT *et al.*, 2023; DUARTE, 2023; COSTA *et al.*, 2023; KLEIN *et al.*, 2023; PALLAVI *et al.*, 2014).

Tendo em vista a importância da seleção de genótipos superiores por meio de multi características de maneira simultânea que atendam as exigência do mercado de soja no Brasil, o objetivo deste trabalho foi comparar diferentes cenários para construção do ideótipo com o uso do MGIDI para a seleção de linhagens superiores de soja.

2. MATERIAIS E MÉTODOS

2.1 Experimento

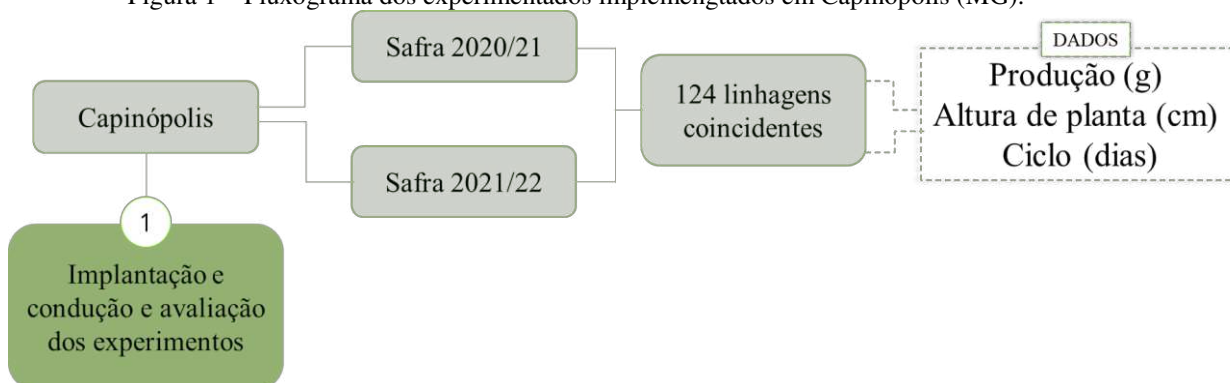
Neste estudo, linhagens de soja foram testadas na localidade de Capinópolis (18°40'48"S, 49°33'58"W, 530 m de altitude) no Estado de Minas Gerais. Todos os ensaios foram instalados em Delinemaneto de Blocos Aumentados e plantados com densidade de semeadura de 12 plantas por metro.

Os ensaios foram conduzidos durante os anos agrícolas de: (1) 2020/21, sendo constituído de 241 genótipos regulares e 48 testemunhas, os quais eram compostos por parcelas de duas linhas de 2,5 m cada, distribuídas em oito blocos, totalizando 289 parcelas e (2) 2021/22, sendo constituído de 128 genótipos regulares e 46 testemunhas, os quais eram compostos por parcelas de uma linha de 3 m, distribuídas em oito blocos, totalizando 174 parcelas. Para ambos os ensaios as testemunhas utilizadas foram: BMX Zeus IPRO (Grau de Maturidade Relativa (GMR) = 5.5), BMX Alvo RR (GMR = 5.9), BS2606 IPRO (GMR = 6.0), TMG 7063 IPRO (GMR = 6.3), DM 66i68 RSF IPRO (GMR= 6.6) e FPS Júpiter RR (GMR= 6.0).

Avaliou-se as seguintes características: produção de grãos por parcela, ciclo e altura da planta na maturação. As parcelas foram colhidas, trilhadas em bulk e as sementes foram pesadas (gramas) para compor o dado de produção por parcela em gramas. A altura de planta na maturação, em centímetros, foi medida quando mais de 50% das plantas na se encontravam no estágio R8, conforme estabelecido por Fehr e Cavinnes (1977). Com

auxílio de uma régua, a planta foi medida da base à extremidade da haste principal. Para o ciclo, estabeleceu o período em dias entre a data de semeadura e a data de maturação, quando mais de 50% das plantas na se encontravam no estágio R8. Os tratos culturais durante o ciclo foram conduzidos segundo as necessidades da cultura da soja para essa região (SEDIYAMA; SILVA; BOREM, 2022).

Figura 1 – Fluxograma dos experimentados implemngtados em Capinópolis (MG).



Fonte: Elaboração própria.

2.2 Análises estatísticas

2.2.1 Máxima Verossimilhança Restrita/Melhor Predição Linear não Viesada (REML/BLUP)

Previamente, os componentes de variância foram obtidos pela metodologia de modelos mistos Máxima Verossimilhança Restrita – REML (PATTERSON & THOMPSON, 1971) e os valores genotípicos foram preditos utilizando Melhor Predição Linear não Viesada – BLUP (HENDERSON, 1975). O método utilizado foi o de análises de séries temporais auto-regressivo de duas dimensões (AR1 x AR1). As testemunhas e os blocos foram tratados como de efeitos fixos e os genótipos regulares como de efeitos aleatórios. O modelo utilizado foi o geral proposto por (Henderson, 1984):

$$y = X\beta + Z\gamma + e \quad (1)$$

Onde,

y é o vetor dos dados, ordenados como colunas e linhas;

β é o vetor paramétrico dos efeitos fixos (blocos e testeunhas);

X é a matriz de incidência dos efeitos contidos em β ;

γ é o vetor paramétrico dos efeitos aleatórios;

Z é a matriz de incidência dos efeitos contidos em γ e

e são os erros que compõem o vetor ($e = \xi + \eta$, onde ξ é o efeito de tendência e η é o resíduo aleatório não correlacionado).

Todos os dados foram analisados por meio do *software* R 2023.03.1 (R Core Team, 2023). Os componentes de variância e os BLUPs foram obtidos utilizando-se o pacote ASReml (Gilmour et al. & Thompson, 2015).

2.2.2 Índice de Distância Genótipo-Ideótipo Multicaracterística (MGIDI)

De posse dos BLUPs de cada característica para cada safra, foi realizada a seleção dos genótipos mais produtivos por meio do Índice de Distância Genótipo-Ideótipo Multicaracterística (MGIDI). O índice foi aplicado considerando a taxa de seleção de 15% e, para a análise, foram considerados somente os genótipos coincidentes entre a safras (Tabela 5), visto que o objetivo é selecionar genótipos para comporem ensaios avançados de avaliação.

Cada característica de cada safra foi considerada uma variável, totalizando seis variáveis, assim tem-se as seguintes variáveis analisadas: produção da safra 2020/21 (PS1), altura de planta da safra 2020/21 (AS1), ciclo da safra 2020/21 (CS1), produção da safra 2021/22 (PS2), altura de planta da safra 2021/22 (AS2), ciclo da safra 2021/22 (CS2).

Na análise de modelos mistos e na aplicação do índice de seleção, todas as variáveis foram consideradas para obtenção dos valores genéticos e dos índices MGIDI. Entretanto, para interpretação dos ganhos genéticos preditos e tomadas de decisão, somente as variáveis da segunda safra foram levadas em consideração, haja visto que a seleção deve ser realizada apenas para a última safra plantada para que os melhores indivíduos sejam selecionados e avançados para a safra seguinte.

O cálculo do índice de distância genótipo-ideótipo multicaracterística seguiu quatro etapas, sendo elas:

i) Nesta primeira etapa foi definida a construção de ideótipos, objetivando maior produção, precocidade e plantas de menor estatura. Foram considerados seis cenários para a construção do ideótipo.

Cenário 1 (C1): Ideótipo construído apenas com base no direcionamento dos sentidos para cada característica (Tabela 1).

Tabela 1 – Sentido para as características e valores de máximo (n_{nj}) e mínimo (φ_{nj}) estabelecidos para as características para o cálculo do índice.

Variável	Sentido	(n_{nj})	(φ_{nj})
PS1	↑	100	0
AS1	↓	0	100
CS1	↓	0	100
PS2	↑	100	0
AS2	↓	0	100
CS2	↓	0	100

PS1 = produção da safra 2020/21 ; AS1 = altura de planta da safra 2020/21 ; CS1 = ciclo da safra 2020/21; PS2 = produção da safra 2021/22; AS2 = altura de planta da safra 2020/22 e CS2 = ciclo de planta da safra 2020/22.

Fonte: Elaboração própria.

Cenário 2 (C2): Ideótipo construído com base no sentido e com base no peso atribuído para cada característica. Os pesos foram estabelecidos por meio da proporção da Correlação de Pearson entre a variável principal, estabelecida como sendo a produção de grãos por parcela na safra 2021/22, e as demais características avaliadas (Tabela 2). Os pesos estabelecidos visam priorizar a característica alvo no processo de seleção e possui como primícias as causas de correlação genética: pleiotropia e genes ligados. Para tanto, calculou-se a correlação de todas as variáveis com a variável PS2, em seguida dividiu-se cada correlação obtida pela correlação total, como descrito abaixo:

$$P_i = r_i / r_t \quad (2)$$

Onde,

P_i é peso atribuída à i -ésima característica inclusive à produção de grãos por parcela na safra 2021/22 ;

r_i é a correlação da i -ésima característica com a produção de grãos por parcela na safra 2021/22, inclusive entre esta última e ela mesma; e,

r_t é soma de todas as correlações.

Tabela 2 – Pesos considerados para cada característica na construção do ideótipo.

Variáveis	Correlação de Pearson	Peso	
PS2	PS1	0,2198	0,1356
	AS1	0,2529	0,1560
	CS1	0,1316	0,0812
	PS2	1,0000	0,6169
	AS2	0,0137	0,0084
	CS2	0,0031	0,0019
Soma	1,62	1,00	

PS1 = produção da safra um ; AS1 = altura de planta da safra um ; CS1 = ciclo da safra um; PS2 = produção da safra dois; AS2 = altura de planta da safra dois e CS2 = ciclo de planta da safra dois.

Fonte: Elaboração própria.

Cenário 3 (C3): Ideótipo construído com base nos valores da média de cada característica da testemunha de melhor performance (Tabela 2).

Cenário 4 (C4): Ideótipo construído com base nos valores da média de cada característica da testemunha de melhor performance e com base em pesos para cada característica (Tabelas 2 e 3).

Tabela 3 – Média dos valores reais para cada característica da melhor testemunha e valores objetivados para uma cultivar desejada.

Ideótipo	Produção (g/parcela)	Produtividade (sc/ ha)	Altura (cm)	Ciclo (dias)
Testemunha	997	110	81	120
Cultivar Alvo	1080	120	80	110

Fonte: Elaboração própria.

Cenário 5 (C5): Ideótipo construído com base nos valores para cada característica de uma cultivar alvo que ultrapasse a performance da testemunha de melhor desempenho (Tabela 3).

Cenário 6 (C6): Ideótipo construído com base nos valores para cada característica de uma cultivar alvo que ultrapasse a performance da testemunha de melhor desempenho e com base em pesos atribuídos às características (Tabelas 2 e 3).

ii) Na segunda etapa a matriz contendo os BLUP's foi reescalada para valores assumam intervalos de 0-100 por meio da equação:

$$rX_{ijn} = \frac{n_{nj} - \varphi_{ij}}{n_{oj} - \varphi_{oj}} \times \theta_{ij} - n_{ij} + n_{nj} \quad (3)$$

Onde,

n_{nj} e φ_{nj} são os valores máximos e mínimos reescalados, respectivamente;
 n_{oj} e φ_{oj} são os valores originais de máximo e mínimo, respectivamente e
 θ_{ij} é o valor original da j -ésima característica no i -ésimo genótipo.

A reescalação dos valores foi feita da seguinte maneira: para características quando um baixo valor foi objetivado $n_{nj} = 0$ e $\varphi_{nj} = 100$ e para características em que um alto valor é objetivado considerou-se, $n_{nj} = 100$ e $\varphi_{nj} = 0$. Deste modo, estabeleceu-se o sentido para a análise das variáveis estudadas (Tabela 1).

iii) Esta terceira etapa consistiu em uma análise fatorial exploratória, em que foi possível encontrar estruturas de correlação e, a partir disso, reduzir a dimensionalidade dos dados. A análise foi realizada seguindo a seguinte equação:

$$F = Z(A^T R^{-1})^T \quad (4)$$

Onde,

F é a matriz ($g \times f$) de escores fatoriais;

Z é a matriz ($g \times p$) com as médias reescaladas;

A é a matriz ($p \times f$) de lodgings canônicas;

R é a matriz ($p \times p$) de correlação entre índices e

g , f e p correspondem, respectivamente, a genótipos, fatores e índices calculados.

iv) Por fim, o MGIDI foi aplicado para calcular as distâncias entre genótipos e ideótipos, conforme estabelecido por Olivoto e Nardino (2021):

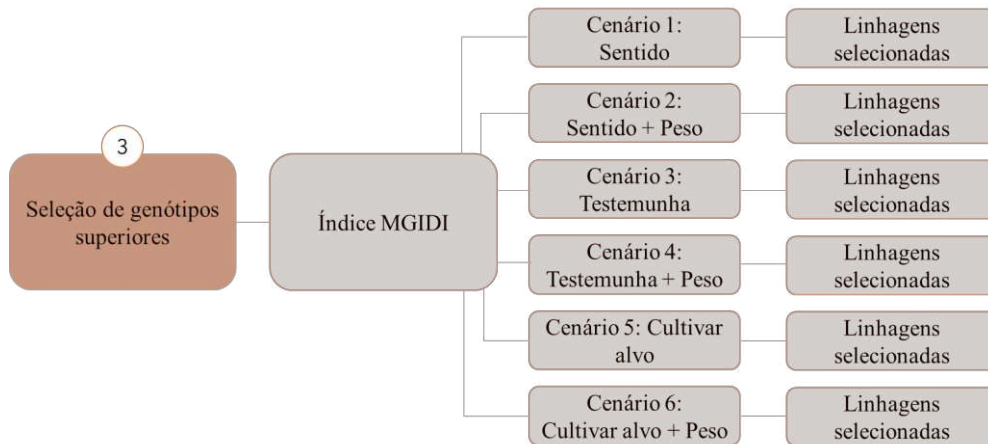
$$MGIDI_i = \sum_{j=1}^f (\gamma_{ij} - \gamma_{j-})^2]^{0.5} \quad (5)$$

Onde,

MGIDI_i: índice da distância entre genótipo-ideótipo;

γ_{ij} : pontuação do i -ésimo genótipo no j -ésimo fator ($i = 1, 2, \dots, g; j = 1, 2, \dots, f$), sendo g e f o número de genótipos e fatores, respectivamente;
 γ_{j-} : o j -ésimo pontuação do ideótipo.

Figura 2 – Fluxograma das análises realizadas nos experimentos de Capinópolis (MG).



Fonte: Elaboração própria.

De posse dos valores do índice de cada genótipo avaliado, os mesmos foram ranqueados e selecionados para cada cenário. Para efeito de comparação dos cenários e escolha da melhor estratégia de seleção, os cenários foram comparados por meios dos ganhos de seleção (GS) obtidos nas análises para cada característica, pela soma dos ganhos de seleção ajustados (GSA) de cada cenário e pelo índice de coincidência dos genótipos selecionados (IC) entre os cenários. O índice de coincidência foi calculado da seguinte maneira:

$$IC_{(\%)} = \frac{nC}{nT_i} \quad (6)$$

Onde,

$IC_{(\%)}$ = Índice de Coincidência;

nC = é o número de genótipos selecionados coincidentes entre os cenários;

nT_i = é o número de genótipos selecionados de acordo com a taxa de seleção de 15%.

A GSA é o GS com sinal ajustado para positivo ou negativo para efeito de soma dos ganhos genéticos e obtenção do ganho total para cada cenário. Sinal positivo se o

ganho foi no sentido requerido e sinal negativo se o ganho foi obtido no sentido contrário do requerido. Veja um exemplo: no C3 a variável CS2 obteve ganho no sentido requerido (negativo), portanto ganhou sinal positivo na soma do GSA, já a variável AS2 obteve ganho no sentido contrário do requerido (positivo), portanto ganhou sinal negativo para a somas da GSA.

Todos os dados foram analisados por meio do *software* R 2023.03.1 (R Core Team, 2023). Os índices MGIDI foram obtidos por meio do pacote Metan (OLIVOTO & NARDINO, 2021).

3. RESULTADOS E DISCUSSÃO

Dentre todos os genótipos analisados por meio do modelo misto, 124 genótipos coincidem para as duas safras. Os 124 indivíduos foram analisados por meio do índice de seleção, dos quais, adotando um índice de seleção de 15%, 19 genótipos superiores foram selecionados para cada cenário.

3.1 Análise de Fatores

Após definição dos ideótipos na primeira etapa foi aplicada a análise fatorial exploratória para todos os cenários. O índice explora apenas os fatores que possuem autovalores (λ) com valor acima de um, tal qual a metodologia de Kaiser (1958). Percebe-se que para os cenários que consideraram apenas o sentido e o peso para as características foram resumidos em três fatores principais. Já os cenários que consideraram valores para as a características foram resumidos em dois fatores.

Na Tabela 1 são demonstrados os resultados da análise de fatores para todos os cenários. Na 4, por meio dos gráficos, é possível verificar as fortalezas e fragilidades de cada genótipo selecionado. Nos gráficos, a contribuição de cada fator para o MGIDI é ranqueado do maior contribuidor (próximo ao centro) para o fator de menor contribuição (próximo à borda do círculo). O círculo pontilhado indica um valor teórico de contribuição igualitária de todos os fatores (OLIVOTO *et al.*, 2022).

Para o cenário 1 (C1) e para o cenário 2 (C2), ambos os conjuntos de dados foram reduzidos para três fatores que, juntos, somam 70,5 % da variância acumulada (σ_A^2). Para o fator um (FA1), as características agrupadas foram ciclo da safra 2020/21 (CS1) e altura de planta da safra 2021/22 (AS2), os quais se correlacionam de forma negativa. As

características que mais se correlacionaram no segundo fator (FA2) foram produção da safra 2020/21 (PS1) e produção da safra 2021/22 (PS2). Já para o fator três (FA3), a variável ciclo da safra 2021/22 (CS2) explica 92% da variância deste fator.

Os conjuntos de dados do cenário três (C3), do cenário 4 (C4) e do cenário 5 (C5) foram reduzidos para dois fatores que explicam 53,20% da σ_A^2 . As características CS1 e AS2 foram as que mais se correlacionaram com o FA1, estas se correlacionaram de forma negativa. Já para o FA2 as características que mais se correlacionaram foram produção da safra 2021/22 (PS2) e ciclo da safra 2021/22 (CS2), estas se correlacionaram de maneira inversa.

3.2 Seleção por MGIDI

Os genótipos selecionados e seus respectivos valores de MGIDI para cada cenário são apresentados na Tabela 7 e na Figura 3. Quanto menor o valor de MGIDI, menor será a distância do genótipo do ideótipo estabelecido, ou seja, mais próximo o genótipo selecionado estará do desejado (OLIVOTO; NARDINO, 2021).

Os cenários foram comparados por meio do ganho de seleção (GS), da soma dos ganhos de seleção ajustados (GSA) e o índice de coincidência (IC) (Tabela 5 e Tabela 6). O GS permitiu avaliar as características individualmente para cada cenário, já a soma dos GSA permitiu avaliar o cenário como um todo, informando o quanto de ganho o cenário proporcionou para todas as características em conjunto. Os valores em verde indicam ganhos no sentido requerido, já os valores em vermelho indicam ganhos obtidos no sentido contrário do requerido. A exemplo disso, no C3 tem-se um GS% no sentido requerido (negativo) para CS2 (GS% = -0,51), já no C2 tem-se um GS% no sentido contrário do requerido (negativo) para PS2 (GS% = - 2,08).

Quando se analisa o GSA% total de cada cenário, a ordem decrescente dos cenários para ganho genético é: C1 (12,14 %) > C5 (9,26) > C3 (8,11) > C4 (-3,64) > C6 (-6,62) > C2 (-8,21), ou seja, o cenário que possibilitou maior ganho genético total foi o cenário 1, o qual considera apenas o sentido de cada característica na construção do ideótipo. É interessante ressaltar que os cenários que obtiveram os piores ganhos genéticos totais foram os que levaram em consideração os pesos atribuídos às características. Vale ressaltar que o C1 foi aquele que apresentou os ganhos mais equilibrados dentre todas as características. Klein *et al.* (2023) relatam que estes ganhos

individuais menores, em alguns casos, pode ser recompensado pelo ganho total simultâneo das características.

Ao analisar a $GS\%$ individual para cada cenário, nota-se que para a variável ciclo, apesar de não muito expressivo, foram obtidos ganhos no C1, C3, C4 e C6 (-0,47; -0,51; -0,67 e -1,39; respectivamente). Em número de dias, esses ganhos representam um dia para C1, C2 e C3 e dois dias para C6. Para a característica altura de planta, a média dos genótipos selecionados em todos os cenários já contem os ideótipos contruídos, todavia, ainda assim, foram obtidos ganhos no C1 (-7,76) e C4 (-1,98).

Embora seja importante considerar o ganho para todas as características desejadas, é importante reiterar que, em um programa de melhoramento genético de soja, o maior objetivo é o incremento da produção de grãos (PRETTI, 2023; BORDIN *et al.*, 2023). Tendo isto em vista, nota-se que a produção de grãos obteve maior incremento nos cenários C1, C3 e C5. Já para os demais cenários, a produção obteve ganho no sentido contrário do esperado. A ordem decrescente de ganho genético para a variável produção foi: C5 (9,95) > C3 (9,57) > C1 (3,91) > C2 (-2,08) > C6 (-6,28) > C4 (-6,28).

Josefh *et al.*, em estudo para seleção de genótipos de soja mais produtivos, encontrou um $GS = 9,18$ quando utilizou o MGIDI e um $GS = 8,97$ quando utilizou o Índice de Mulamba e Mock (1978), o que corrobora com os dados encontrados neste estudo. Note que, analisando os $GS\%$ individuais, também foram obtidos ganhos no sentido desejado apenas para os cenários que não levaram em consideração os pesos, o que reitera os resultados obtidos pela análise de GSA. Os maiores ganhos foram obtidos nos cenários em que os ideótipos foram definidos pelos valores da melhor testemunha e pelos valores da cultivar alvo, sendo o melhor cenário, em termos de ganho de seleção, o que considerou valores de uma cultivar desejada para a construção do ideótipo (C5).

Tabela 4 – Cargas fatoriais para cada variável e para cada cenário.

Variáveis	Cenário 1			Cenário 2			Cenário 3		Cenário 4		Cenário 5		Cenário 6	
	FA1	FA2	FA3	FA1	FA2	FA3	FA1	FA2	FA1	FA2	FA1	FA2	FA1	FA2
PS1	0,33	-0,72	-0,18	0,33	-0,72	-0,18	-0,50	-0,12	-0,50	-0,12	-0,50	-0,19	-0,50	-0,19
AS1	0,43	-0,32	0,16	0,43	-0,32	0,16	-0,49	-0,05	-0,49	-0,05	-0,48	-0,12	-0,48	-0,12
CS1	-0,90	0,04	-0,03	-0,90	0,04	-0,03	-0,91	-0,03	-0,91	-0,03	-0,91	0,02	-0,91	0,02
PS2	-0,34	-0,77	0,15	-0,34	-0,77	0,15	-0,18	-0,73	-0,18	-0,73	-0,14	-0,78	-0,14	-0,78
AS2	-0,88	-0,10	0,08	-0,88	-0,10	0,08	-0,83	0,05	-0,83	0,05	0,83	-0,15	0,83	-0,15
CS2	0,03	0,00	0,97	0,03	0,00	0,97	-0,07	0,78	-0,07	0,78	-0,03	0,72	-0,03	0,72
	1,99	1,23	1,01	1,99	1,23	1,01	2,05	1,14	2,05	1,14	2,01	1,19	2,01	1,19
	33,20	20,40	16,80	33,20	20,40	16,80	34,10	19,10	34,10	19,10	33,60	19,80	33,60	19,80
	33,20	53,70	70,50	33,20	53,70	70,50	34,10	53,20	34,10	53,20	33,60	53,30	33,60	53,30

PS1 = produção da safra um ; AS1 = altura de planta da safra um ; CS1 = ciclo da safra um; PS2 = produção da safra dois; AS2 = altura de planta da safra dois e CS2 = ciclo de planta da safra dois; FA1 = fator um; FA2 = fator 2 e FA3= fator 3.

Fonte: Elaboração própria.

Tabela 5 – Ganhos genéticos para os materiais selecionados.

Cenário 1					
Variáveis	\bar{X}_o	\bar{X}_s	GS	GS (%)	GSA (%)
PS1	628,00	628,00	0,00	0,00	0,00
AS1	70,50	66,90	-3,60	-5,12	5,12
CS1	112,00	112,00	-0,08	-0,08	0,08
PS2	695,00	723,00	27,20	3,91	3,91
AS2	82,20	75,80	-6,37	-7,76	7,76
CS2	131,00	130,00	-0,62	-0,47	0,47
Total					17,34

Cenário 2					
Variáveis	\bar{X}_o	\bar{X}_s	GS	GS (%)	GSA (%)
PS1	628,00	628,00	0,00	0,00	0,00
AS1	70,50	77,30	6,80	9,65	-9,65
CS1	112,00	112,00	0,15	0,14	-0,14
PS2	695,00	681,00	-14,50	-2,08	-2,08
AS2	82,20	87,00	4,84	5,90	-5,90
CS2	131,00	131,00	0,31	0,23	-0,23
Total					-18,00

Cenário 3					
Variáveis	\bar{X}_o	\bar{X}_s	GS	GS (%)	GSA (%)
PS1	628,00	628,00	0,00	0,00	0,00
AS1	70,50	80,20	9,69	13,80	-13,80
CS1	112	112	0,168	0,15	-0,15
PS2	695	762	66,6	9,57	9,57
AS2	82,2	83,8	1,62	1,97	-1,97
CS2	131	130	-0,665	-0,508	0,508
Total					-5,84

Cenário 4					
Variáveis	\bar{X}_o	\bar{X}_s	GS	GS (%)	GSA (%)
PS1	628,00	628,00	0,00	0,00	0,00
AS1	70,50	69,90	-0,60	-0,01	0,01
CS1	112,00	112,00	-0,09	-0,08	0,08
PS2	695,00	652,00	-43,70	-6,28	-6,28
AS2	82,20	83,80	1,62	-1,98	1,98
CS2	131,00	130,00	-0,87	-0,67	0,67
Total					-3,54

Cenário 5					
Variáveis	\bar{X}_o	\bar{X}_s	GS	GS (%)	GSA (%)
PS1	628,00	628,00	0,00	0,00	0,00
AS1	70,50	79,70	9,20	13,10	-13,10
CS1	112,00	112,00	0,16	0,14	-0,14
PS2	695,00	765,00	69,20	9,95	9,95
AS2	82,20	82,60	0,41	0,50	-0,50
CS2	131,00	131,00	0,24	0,19	-0,19
Total					-3,98

Cenário 6					
Variáveis	\bar{X}_o	\bar{X}_s	GS	GS (%)	GSA (%)
PS1	628,00	628,00	0,00	0,00	0,00
AS1	70,50	69,70	-0,77	-1,09	1,09
CS1	112,00	112,00	0,00	0,00	0,00
PS2	695,00	655,00	-40,50	-5,82	-5,82
AS2	82,20	84,00	1,80	2,19	-2,19
CS2	131,00	129,00	-1,82	-1,39	1,39
Total					-5,58

PS1 = produção da safra um ; AS1 = altura de planta da safra um ; CS1 = ciclo da safra um; PS2 = produção da safra dois; AS2 = altura de planta da safra dois e CS2 = ciclo de planta da safra dois; \bar{X}_o = média da população original; \bar{X}_s = média da população selecionada; GS = ganho de seleção; GS % = percentual do ganho de seleção; GSA% = ganho de seleção ajustado

Fonte: Elaboração própria.

O ranking dos genótipos selecionados estão representados graficamente na Figura 1. O círculo em vermelho aponta a intensidade de seleção aplicada no índice, ou seja, o ponto de corte aplicado para a seleção (POUR-ABOUGHADAREH; POCZAI, 2021). Os genótipos posicionados externamente ao círculo foram selecionados (sinalizados em pontos vermelhos) e os genótipos localizados dentro do círculo não foram selecionados (sinalizados em pontos cinzas). Por meio dos Índices de Coincidência (IC%), é possível verificar a relação dos genótipos selecionados entre os cenários (Tabelas 6 e 7).

Tabela 6 – Ranking dos Índices de Coincidências (IC%) entre os genótipos selecionados nos diferentes cenários analisados.

	Cenários	IC
C3+C5	Testemunha + Cultivar alvo	74%
C4+C6	Testemunha com peso + Cultivar alvo com peso	63%
C2+C3	Sentido com peso + Testemunha	37%
C2+C5	Sentido com peso + Cultivar alvo	37%
C2+C6	Sentido com peso + Cultivar alvo com peso	32%
C1+C4	Sentido + Testemunha com peso	26%
C1+C6	Sentido + Cultivar alvo com peso	26%
C3+C6	Testemunha + Cultivar alvo com peso	26%
C2+C4	Sentido com peso + Testemunha com peso	16%
C1+ C3	Sentido + Testemunha	11%
C1+C5	Sentido + Cultivar alvo	5%
C1 + C2	Sentido + Sentido com Peso	0%
C3+C4	Testemunha + Testemunha com peso	0%
C4+C5	Testemunha com peso + Cultivar alvo	0%
C5+C6	Cultivar alvo + Cultivar alvo com peso	0%

Fonte: Elaboração própria.

O maior IC ocorreu para os C3 e C5 (14 genótipos coincidentes), os quais definem o ideótipo por meio dos valores da testemunha e da cultivar alvo, respectivamente. O segundo maior IC aconteceu para C4 e C6 (12 genótipos coincidentes), os quais consideram, respectivamente, testemunha com peso + cultivar alvo com peso para a construção do ideótipo. Em terceiro, C2 e C3 que consideram sentido com peso + testemunha e também C2 e e C5 que consideram sentido com peso + cultivar alvo com (7 genótipos coincidentes). Em seguida vieram: C2 e C6 (6 genótipos coincidentes), C1 e C4 (5 genótipos coincidentes), C1 e C6 (5 genótipos coincidentes), C3 + C6 (5 genótipos coincidentes), C2 e C4 (3 genótipos coincidentes), C1 e C3 (2 genótipos coincidentes), C1 e C5 (1 genótipo coincidente), C1 e C2 (0 genótipo coincidentes), C3 + C4 (0 genótipo coincidentes), C4 e C5 (0 genótipo coincidentes).

Conforme esperado, de acordo com os resultados anteriormente obtidos, as maiores coincidências acontecem entre as comparações realizadas entre os ideótipos construídos considerando-se testemunha e cultivar desejada. No entanto, ao contrário do esperado, quando ideótipos com testemunha (C3) e cultivar alvo (C5) são comparados com ideótipos sentido com peso (C2), as coincidências são maiores do que quando comparadas somente com sentido (C1). Isto pode ter ocorrido em virtude do número de características que os cenários obtiveram ganhos no sentido não desejado coincidirem mais quando são comparados com o C2 do que com o C1.

A linhagem elite que mais se destacou foi a MDC1312S7, esta ficou em primeiro lugar nos C3 e C5 e em sétimo lugar no C1, os quais foram os cenários que obtiveram os maiores ganhos para produção. Tais resultados sugerem que utilizar somente valores para a construção do ideótipo foi mais assertivo para a obtenção de maiores ganhos genéticos para este conjunto de dados. Todavia, ressalta-se que a utilização dos valores da testemunha proporcionou ganhos, porém foram selecionados genótipos que mais se aproximam da testemunha, sem considerar que os valores da testemunha podem ser superados.

Já o ideótipo que é construído com valores desejados de uma cultivar para a região, com valores acima da média da melhor testemunha, os indivíduos selecionados ultrapassam a performance da testemunha, fazendo assim, com que a média da população selecionada seja maior que a média da melhor testemunha. Os resultados também sugerem que o peso e/ou o peso denotado por meio da correlação de Pearson pode não ter sido uma boa estratégia para a seleção de genótipos superiores. Isto pode ter ocorrido em vista da pouca variância para a característica CS2 e também em virtude de os mesmos pesos terem sido considerados para as duas safras. Visto que a interação genótipo ambiente é diferente e muito forte para cada ambiente, construir pesos para cada safra separadamente pode ser mais eficiente.

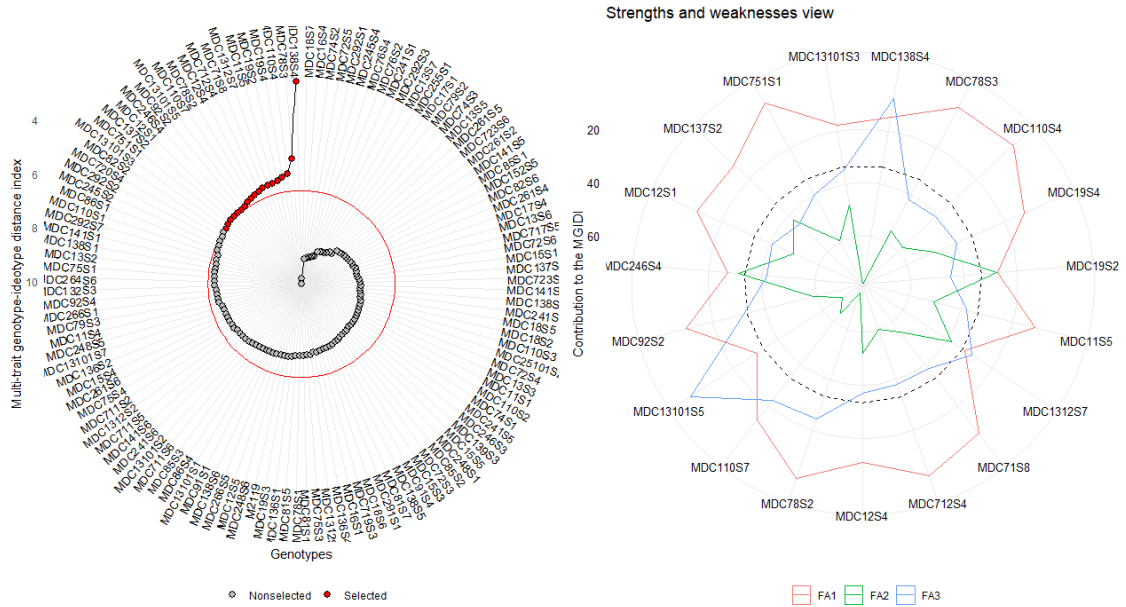
Tabela 7 – Ranking dos genótipos selecionados e seus respectivos índices MGIDI.

Ranking	Cenário 1		Cenário 2		Cenário 3		Cenário 4		Cenário 5		Cenário 6	
	Genótipo	MGIDI	Genótipo	MGIDI	Genótipo	MGIDI	Genótipo	MGIDI	Genótipo	MGIDI	Genótipo	MGIDI
1	MDC138S4	2,58	MDC18S7	0,42	MDC1312S7	2,22	MDC78S3	1,07	MDC1312S7	0,21	MDC138S5	0,15
2	MDC78S3	5,42	MDC16S4	0,65	MDC92S4	2,29	MDC110S4	1,34	MDC139S3	0,22	MDC138S2	0,53
3	MDC110S4	5,99	MDC292S1	1,38	MDC18S7	2,50	MDC74S2	1,81	MDC25101S2	0,51	MDC16S4	1,01
4	MDC19S4	6,08	MDC292S3	1,59	MDC139S3	2,61	MDC241S1	2,04	MDC110S2	0,57	MDC711S5	1,07
5	MDC19S2	6,16	MDC255S1	1,73	MDC13101S5	2,62	MDC138S5	2,06	MDC292S7	0,60	MDC138S1	1,11
6	MDC11S5	6,20	MDC17S1	1,98	MDC137S3	2,83	MDC720S4	2,21	MDC137S3	0,68	MDC245S4	1,17
7	MDC1312S7	6,23	MDC245S4	2,06	MDC292S7	2,84	MDC78S2	2,27	MDC141S5	0,77	MDC13S7	1,40
8	MDC71S8	6,26	MDC141S5	2,14	MDC110S2	2,96	MDC13S7	2,32	MDC241S5	0,84	MDC78S3	1,60
9	MDC712S4	6,32	MDC137S3	2,17	MDC25101S2	2,99	MDC71S8	2,32	MDC15S5	0,86	MDC13101S5	1,69
10	MDC12S4	6,36	MDC13S7	2,34	MDC13S2	3,03	MDC76S4	2,32	MDC246S3	0,87	MDC18S7	1,77
11	MDC78S2	6,41	MDC76S2	2,39	MDC138S1	3,23	MDC138S2	2,34	MDC13S2	0,89	MDC723S6	1,78
12	MDC110S7	6,44	MDC25101S2	2,40	MDC16S4	3,25	MDC245S4	2,37	MDC241S6	0,90	MDC110S4	1,89
13	MDC13101S5	6,54	MDC13S5	2,42	MDC82S2	3,31	MDC82S6	2,49	MDC82S2	0,93	MDC92S4	1,90
14	MDC92S2	6,55	MDC246S3	2,50	MDC241S5	3,36	MDC723S6	2,54	MDC248S1	0,94	MDC138S4	2,06
15	MDC246S4	6,56	MDC261S5	2,50	MDC246S3	3,36	MDC711S5	2,60	MDC141S6	0,96	MDC74S2	2,18
16	MDC12S1	6,56	MDC241S5	2,56	MDC241S6	3,37	MDC110S7	2,62	MDC141S2	0,98	MDC76S2	2,27
17	MDC137S2	6,57	MDC91S4	2,62	MDC15S5	3,38	MDC85S1	2,64	MDC81S5	1,05	MDC17S4	2,27
18	MDC751S1	6,59	MDC723S6	2,62	MDC141S5	3,40	MDC261S4	2,65	MDC91S4	1,09	MDC71S8	2,29
19	MDC13101S3	6,62	MDC13S6	2,65	MDC141S6	3,42	MDC17S4	2,65	MDC255S1	1,12	MDC241S1	2,30
Total		116,44		39,12		56,97		42,66		15,00		30,43

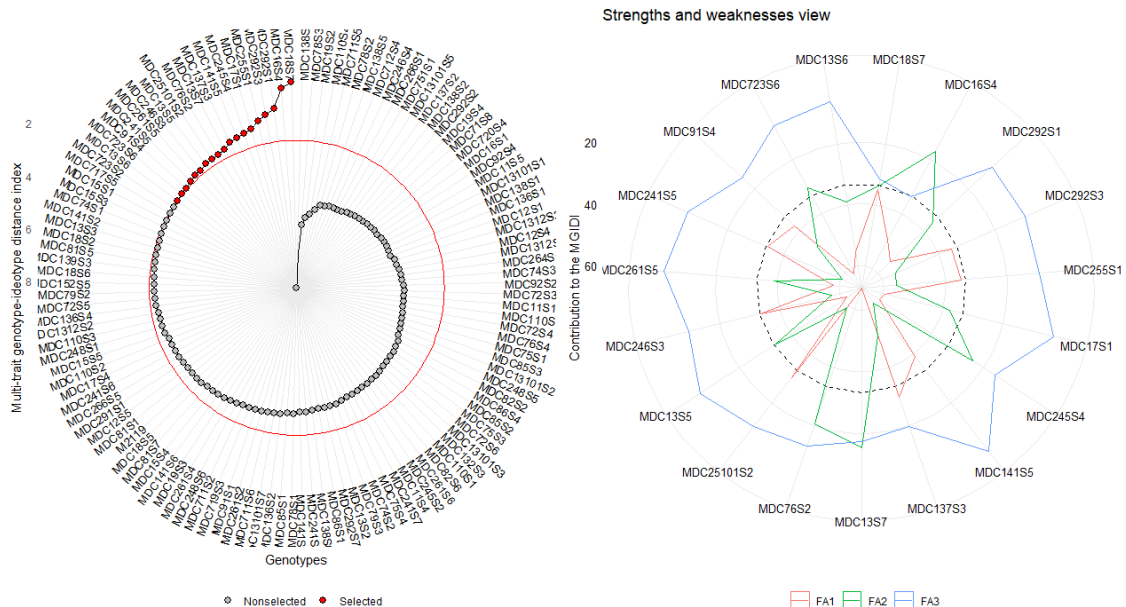
Fonte: Elaboração própria.

Figura 3 – Gráficos do ranking e gráficos das forças e fargilidades dos genótipos selecionados para o cenário 1 (a), cenário 2 (b), cenário 3 (c), cenário 4 (d), cenário 5 (e) e cenário 6 (f).

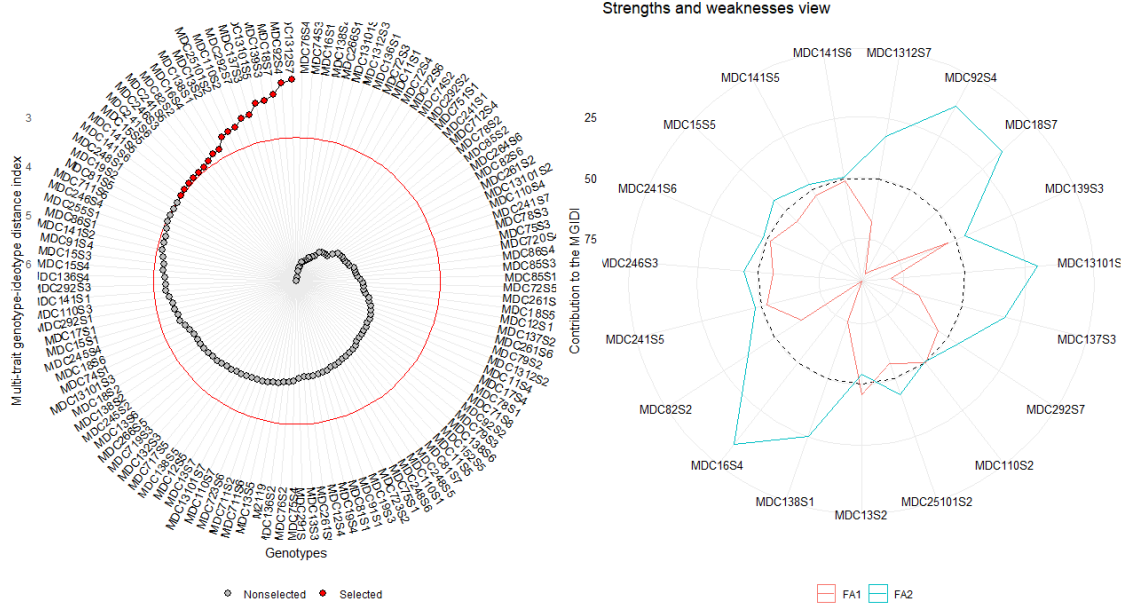
(a)



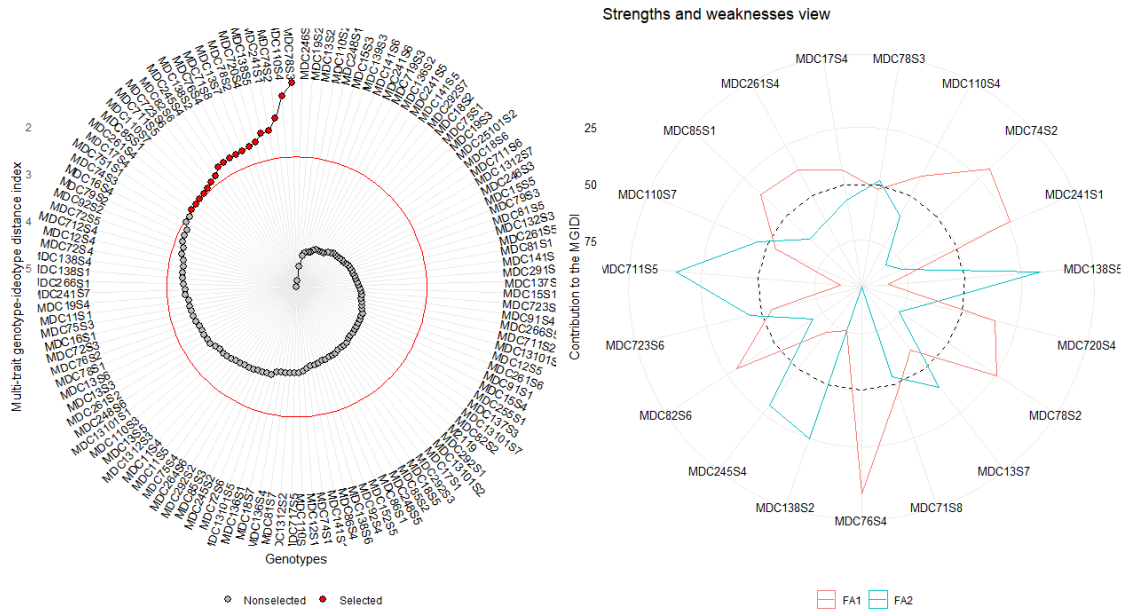
(b)



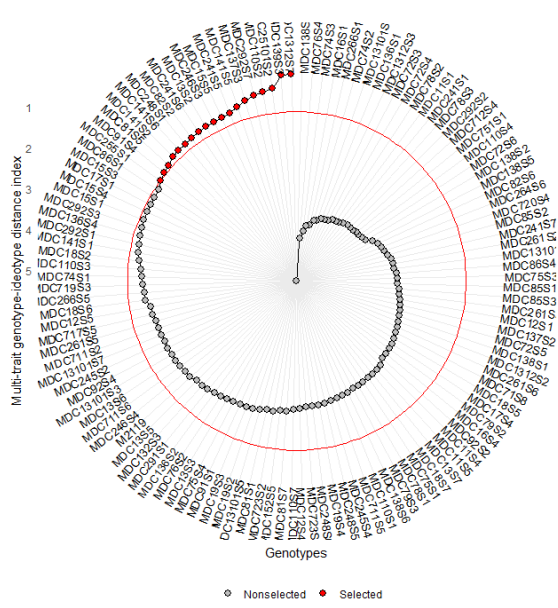
(c)



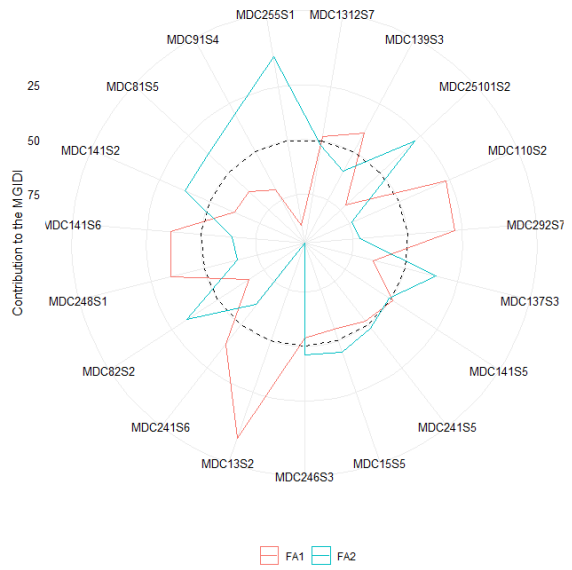
(d)



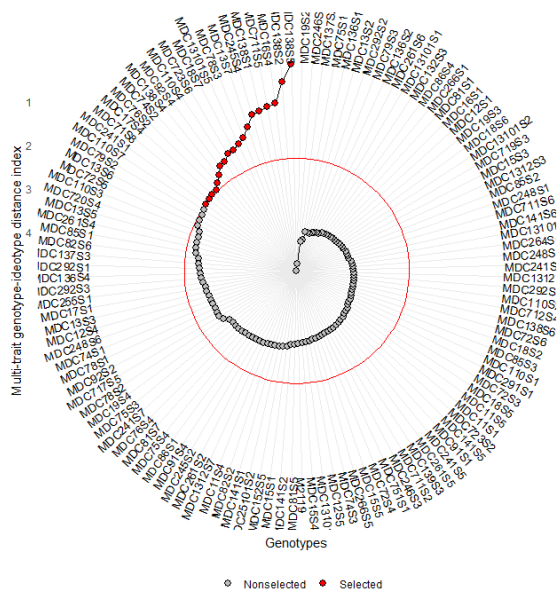
(e)



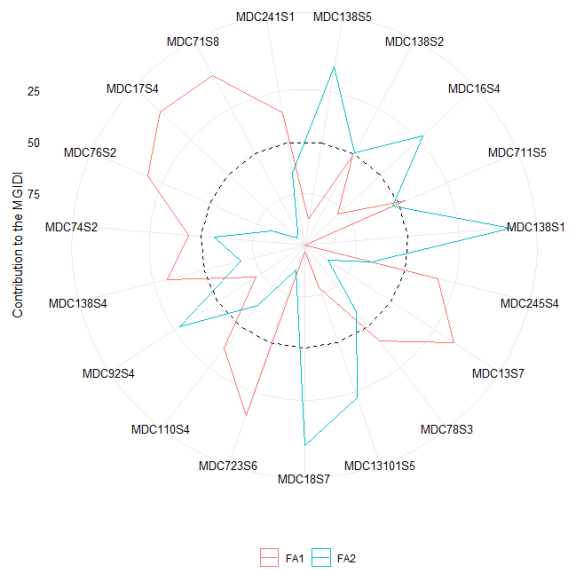
Strengths and weaknesses view



(f)



Strengths and weaknesses view



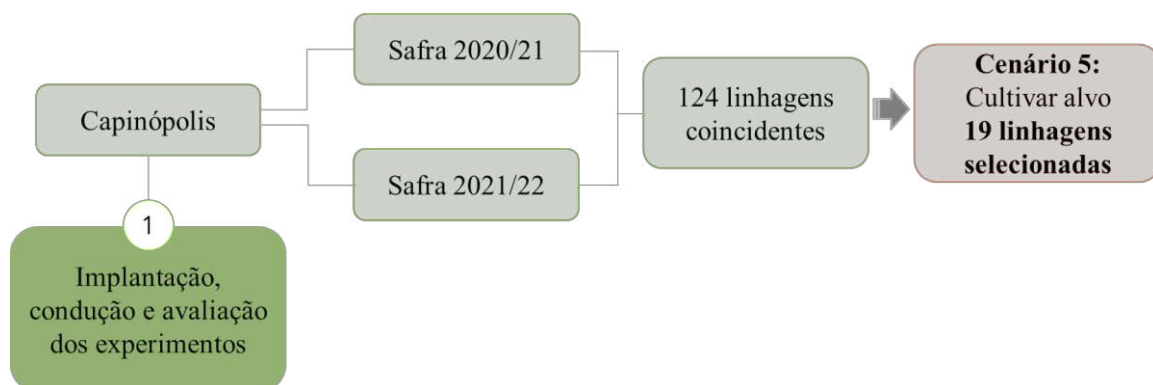
Fonte: Elaboração própria.

4. CONCLUSÕES

Para este conjunto de dados e para estas características, a melhor construção do ideótipo foi a utilizada no cenário cinco (Figura 4), o qual considerou os valores de uma cultivar alvo para as características produção, altura de planta e ciclo.

As linhagens elite do Programa Soja UFV que se destacam como potenciais cultivares para a região de Capinópolis são, em ordem decrescente de proximidade ao ideótipo: MDC1312S7, MDC139S3, MDC25101S2, MDC110S2, MDC292S7, MDC137S3, MDC141S5, MDC241S5, MDC15S5, MDC246S3, MDC13S2, MDC241S6, MDC82S2, MDC248S1, MDC141S6, MDC141S2, MDC81S5, MDC91S4 e MDC255S1.

Figura 4 - Fluxograma das conclusões apresentadas para os ensaios realizados em Capinópolis (MG).



Fonte: Elaboração própria.

REFERÊNCIAS

BENAKANAHALLI, Niranjana Kumara *et al.* A Framework for Identification of Stable Genotypes Based on MTSI and MGDII Indexes: An Example in Guar (*Cymopsis tetragonoloba* L.). **Agronomy**, v. 11, n. 6, p. 1221, 2021.

BORDIN, Pedro Augusto Nicolete. Desempenho agrônomo e caracterização de genótipos de soja em ensaio preliminar. 2023.

COSTA, Ana Lucia Aranha *et al.* Selection of superior soybean genotypes based on multi-trait using MGIDI and FAI-BLUP indices in multi-environmental trials. 2023.

CRUZ, C. D.; CARNEIRO, P. C. S.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. 5.ed. Viçosa, MG: UFV, 2014. v.1, p.514.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos Biométricos Aplicados ao Melhoramento Genético** - ISBN 9788572694339. 4. ed. Viçosa, MG: Editora UFV, 2012. v. 1. 514p.

DALAROSA, Leandro Escobar *et al.* **Parâmetros genéticos e índice de seleção MGIDI na identificação de genótipos superiores de mandioca**. 2021.

DE OLIVEIRA, Rebeca Lourenco *et al.* Multitrait selection of pumpkin genotypes aimed at reducing the growth habit and improving seed production. **Crop Science**, v. 61, n. 3, p. 1620-1629, 2021.

DUARTE, Anunciene Barbosa. **Estratégias e seleção de cultivares de soja em estádios iniciais para tolerância ao déficit hídrico**. 2023.

FERREIRA, João Marcos Soares, M.Sc., Universidade Federal de Viçosa, março de 2020. **Índices de seleção baseados em valores genotípicos aplicados ao melhoramento da soja para aumento do teor de proteína**.

GASTL FILHO, Josef *et al.* Genetic parameters and selection strategies for soybean progenies aiming at precocity and grain productivity. **Ciência e Agrotecnologia**, v. 46, p. e004322, 2022.

HAZEL, L.N. The genetic basis for constructing selection indexes. **Genetics**, v.28, p.476-490, 1943.

KLEIN, Luís Antônio *et al.* Selection of superior black oat lines using the MGIDI index. **Crop Breeding and Applied Biotechnology**, 23(3): e45112332, 2023.

MEIER, Carine *et al.* Genetic parameters and multiple-trait selection in wheat genotypes. **Pesquisa Agropecuária Tropical**, v. 51, 2021.

MOREIRA, Adriano Abreu. **Utilização de imagens na seleção de genótipos de soja resistentes ao complexo de percevejos e com elevada qualidade fisiológica de sementes**. 2022. Tese de Doutorado. Universidade de São Paulo.

NASCIMENTO, Heloisa Rocha do, D.Sc., **Identificação de caracteres e seleção de genótipos de soja tolerantes ao deficit hídrico**. 2020.

OLIVEIRA, Isadora Cristina Martins *et al.* Combining ability of biomass sorghum lines for agroindustrial characters and multitrait selection of photosensitive hybrids for energy cogeneration. **Crop Science**, v. 59, n. 4, p. 1554-1566, 2019.

OLIVOTO, Tiago *et al.* MGIDI: a powerful tool to analyze plant multivariate data. **Plant Methods**, v. 18, n. 1, p. 121, 2022.

OLIVOTO, Tiago *et al.* Multivariate analysis of strawberry experiments: where are we now and where can we go?. **BioRxiv**, p. 2020.12. 30.424876, 2021.

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

OLIVOTO, Tiago; NARDINO, Maicon. MGIDI: toward an effective multivariate selection in biological experiments. **Bioinformatics**, v. 37, n. 10, p. 1383-1389, 2021. 41.

PALLAVI, M. *et al.* Multi trait genotype- ideotype distance index (MGIDI) for early seedling vigour and yield related traits to identify elite lines in rice (*Oryza sativa* L.). **Electronic Journal of Plant Breeding**. Vol 15(1) : 120 – 131

POUR-ABOUGHADAREH, Alireza *et al.* MGIDI and WAASB indices: The useful approaches for selection of salt-tolerant barley genotype at the early growth and maturity stages. 2021.

POUR-ABOUGHADAREH, Alireza; POCZAI, Peter. Dataset on the use of MGIDI index in screening drought-tolerant wild wheat accessions at the early growth stage. **Data in Brief**, v. 36, p. 107096, 2021.

POUR-ABOUGHADAREH, Alizera *et al.* Identification of salt-tolerant barley genotypes using multiple-traits index and yield performance at the early growth and maturity stages. **Bulletin of the National Research Centre** (2021) 45:117

PRETTI, Daniela Porto da Mota *et al.* **Caracterização fenotípica e diversidade genética em soja transgênica**. 2023.

ROCHA, J. R. A. S.; MACHADO, J. C.; CARNEIRO, P. C. S. Multitrait index based on factor analysis and ideotype-design: proposal and application on elephant grass breeding for bioenergy. **GCB Bioenergy**. v. 95, n.1, p. 27-32, 2017.

ROCHA, João Romero do Amaral Santos de Carvalho *et al.* Selection of superior inbred progenies toward the common bean ideotype. **Agronomy Journal**, v. 111, n. 3, p. 181-1189, 2019.

SILVA, M. J.; CARNEIRO, P. C. S.; CARNEIRO, J. E.; DAMASCENO, C. M. B.; PARRELLA, N. N. L. D.; PASTINA, M. M.; SIMEONE, M. L. F.; SCHAFFERT, R.

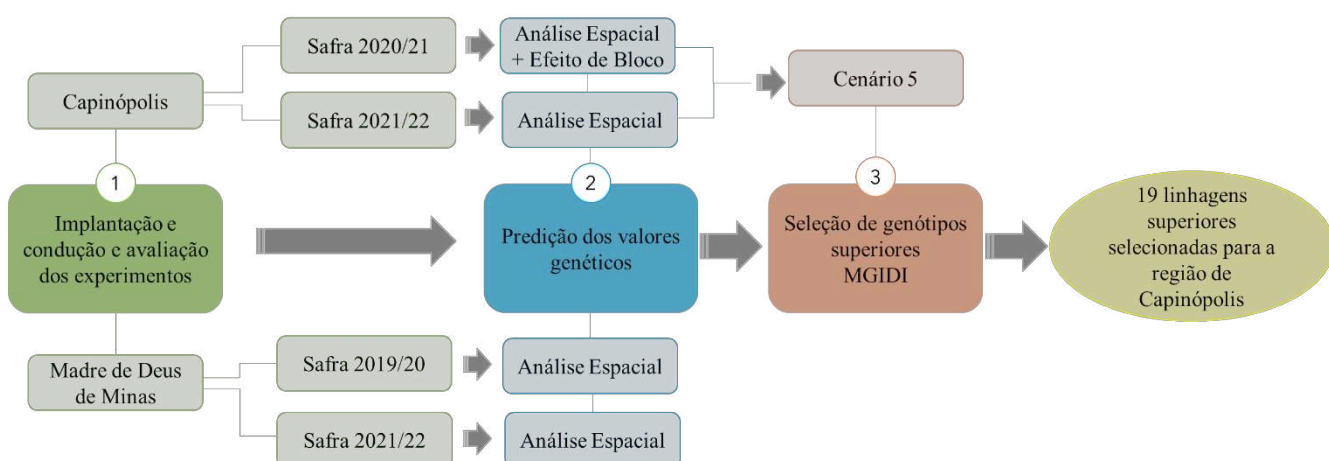
- E.; PARRELLA, R. A. Evaluation of the potential of lines and hybrids of biomass sorghum. **Industrial Crops and Products**, v. 125, n. 1, p. 379-385, 2018
- SMITH, H.F. A discriminant function for plant selection. **Annual Eugenics**, v.7, p.240-250, 1936.
- SOARES, Bruno de Almeida. **Identificação da base genética da resistência parcial de cultivares de soja à *Sclerotinia sclerotiorum***. 2022.
- TISOTT, Jean Vitor *et al.* Agronomic Performance and Selection of Ideal Soybean Genotypes. **Agri-Environmental Sciences**, v. 9, n. 1, p. 12-12, 2023.
- VOLPATO, Leonardo *et al.* Inference of population effect and progeny selection via a multi-trait index in soybean breeding. **Acta Scientiarum. Agronomy**, v. 43, 2020.
- WOYANN, Leomar Guilherme *et al.* Multiple-trait selection of soybean for biodiesel production in Brazil. **Industrial Crops and Products**, v. 140, p. 111721, 2019.

CONCLUSÃO GERAL

Os modelos que fazem uso da análise espacial se ajustaram melhor para os conjuntos de dados analisados.

Para o conjunto de dados de Capinópolis, a construção de ideótipo que mais obtem ganho de seleção para produção é a que considera os valores de uma cultivar alvo.

Figura 5 – Fluxograma da conclusão geral dos ensaios e análises realizados.



Fonte: Elaboração própria.