

**UNIVERSIDADE FEDERAL DE VIÇOSA**

**Informações espaciais incorporadas à análise de vizinhança para avaliação de famílias de cana-de-açúcar**

Rodrigo da Cruz Nunes  
*Magister Scientiae*

**VIÇOSA - MINAS GERAIS  
2025**

**RODRIGO DA CRUZ NUNES**

**Informações espaciais incorporadas à análise de vizinhança para avaliação de famílias de cana-de-açúcar**

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

Orientador: Antonio P. Souza Carneiro

Coorientadores: Luiz A. Peternelli  
Matheus de P. Ferreira

**VIÇOSA - MINAS GERAIS  
2025**

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

T

Nunes, Rodrigo da Cruz, 1998-  
N972i Informações espaciais incorporadas à análise de vizinhança  
2025 para avaliação de famílias de cana-de-açúcar / Rodrigo da Cruz  
Nunes. – Viçosa, MG, 2025.  
1 dissertação eletrônica (51 f.): il. (algumas color.).

Orientador: Antônio Policarpo Souza Carneiro.  
Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Estatística, 2025.

Referências bibliográficas: f. 48-51.

DOI: <https://doi.org/10.47328/ufvbbt.2026.330>

Modo de acesso: World Wide Web.

1. Cana-de-açúcar - Melhoramento genético - Métodos estatísticos. 2. Análise de variância. 3. Análise espacial (Estatística). 4. Planejamento experimental. I. Carneiro, Antônio Policarpo Souza, 1973-. II. Universidade Federal de Viçosa. Departamento de Estatística. Programa de Pós-Graduação em Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 633.612

**RODRIGO DA CRUZ NUNES**

**Informações espaciais incorporadas à análise de vizinhança para avaliação de famílias de cana-de-açúcar**

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

APROVADA: 26 de fevereiro de 2025.

Assentimento:

---

Rodrigo da Cruz Nunes  
Autor

---

Antonio Policarpo Souza Carneiro  
Orientador

Essa dissertação foi assinada digitalmente pelo autor em 02/06/2026 às 13:28:03 e pelo orientador em 02/06/2026 às 17:34:50. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **55R1.IGEF.9EZ1** e clique no botão 'Validar documento'.

À minha mãe, irmãos e sobrinhos, por quem tenho amor incondicional.

## AGRADECIMENTOS

À minha mãe Elizabete e irmãos, Rosângela e Ruan, pelo apoio.

Aos meus amigos do alojamento pós 2022, com quem compartilhei a vida durante a graduação.

À família 218, pelo carinho, incentivo e conselhos.

À minha amiga Mariana, por todo apoio e ajuda nos scripts do R.

Aos meus amigos da Matemática e Pós-graduação, pelas palavras de incentivo, pelo companheirismo e pelos momentos de confraternização.

À Casa Amarela, que em vários momentos foi o meu refúgio.

À Universidade Federal de Viçosa (UFV), pelos auxílios estudantis durante a graduação, fundamentais para minha permanência, e ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, pela oportunidade.

À Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (RIDESA), pela disponibilidade dos dados que viabilizaram a realização desta pesquisa.

Ao meu orientador, Prof. Dr. Antônio Policarpo Souza Carneiro, pela compreensão, disponibilidade e auxílio fundamental durante esses dois anos.

Ao meu coorientador, Matheus de Paula Ferreira, pelo incentivo e ensinamentos.

Ao meu coorientador, Prof. Dr. Luiz Alexandre Peternelli, pelo auxílio e incentivo.

Aos canais do Youtube que disponibilizam conteúdo de qualidade sobre o R e geoestatística, em particular aos canais Bruno Vilela e Canal da Quaest e Geo RGB. Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## RESUMO

NUNES, Rodrigo da Cruz, M.Sc., Universidade Federal de Viçosa, fevereiro de 2025. **Informações espaciais incorporadas à análise de vizinhança para avaliação de famílias de cana-de-açúcar.** Orientador: Antonio Policarpo Souza Carneiro. Coorientadores: Luiz Alexandre Peternelli e Matheus de Paula Ferreira.

No delineamento em blocos casualizados (DBC), os princípios básicos da experimentação repetição, casualização e controle local são fundamentais para reduzir a estimativa do erro experimental. Contudo, em ensaios de avaliação de famílias de cana-de-açúcar, a pouca disponibilidade de material genético e o grande número de genótipos resultam em parcelas pequenas e blocos excessivamente grandes, podendo dificultar a homogeneidade dentro dos blocos e diminuir a eficiência do controle. Uma alternativa para resolver esta questão é a utilização de métodos de vizinhança, como médias móveis e o procedimento de Papadakis, que incorporam a dependência espacial por meio da análise de covariância. Este estudo teve por objetivo avaliar uma abordagem para essas metodologias que utiliza a distância e a covariância espacial como ponderadores. Para isto, foram utilizados dados de dois experimentos de famílias de cana-de-açúcar da RIDESA. O diferencial da metodologia consistiu na utilização do semivariograma para identificar a estrutura de dependência espacial e definir, por meio do alcance efetivo, o número das parcelas vizinhas para o modelo. O experimento 1 apresentou dependência espacial moderada e nele a metodologia proposta superou as análises tradicionais, reduzindo o Coeficiente de Variação de 19,70% para 18,54%, elevando a acurácia seletiva (ASE) de 0,8536 para 0,8716 e tendo uma eficiência relativa de 88,73% na variância do resíduo em relação ao DBC. Essa melhoria na precisão resultou em alterações pontuais no ranqueamento das famílias, evidenciando a capacidade do método em evitar descartes de genótipos promissores. No experimento 2, o DBC foi eficiente em garantir a independência dos resíduos, superando os ajustes de vizinhança em termos de precisão. Concluiu-se que o uso de métodos de vizinhança, em conjunto com o estudo da estrutura espacial, é uma estratégia útil para resolver a baixa eficiência do controle local, se necessário.

Palavras-chave: melhoramento genético; dependência espacial; covariância espacial; precisão experimental; médias móveis; Papadakis

## ABSTRACT

NUNES, Rodrigo da Cruz, M.Sc., Universidade Federal de Viçosa, February, 2025. **Spatial information incorporated into neighborhood analysis for the evaluation of sugarcane families.** Adviser: Antonio Policarpo Souza Carneiro. Co-advisers: Luiz Alexandre Peternelli and Matheus de Paula Ferreira.

In the randomized complete block design (RCBD), the basic principles of experimentation replication, randomization and local control are key to reducing experimental error estimates. However, in sugarcane family evaluation trials, limited genetic material and the high number of genotypes result in small plots and excessively large blocks, which hinders within-block homogeneity and may reduce control efficiency. An alternative to mitigate this issue is the use of neighbor methods, such as moving averages and Papadakis procedure, which capture spatial dependence through covariance analysis. This study aimed to evaluate an approach for these methodologies that utilizes special distance and covariance between plots as weights. Data from two sugarcane family field experiments from RIDESA were used. The methodological highlight consisted of using the semivariogram to identify the spatial dependence structure and define, through the effective range, the number of neighboring plots for the model. Experiment 1 showed moderate spatial dependence, in which the proposed methodology outperformed traditional analyses, reducing the Coefficient of Variation from 19.70% to 18.52%, increasing selective accuracy (SA) from 0.8536 to 0.8716, while achieving a relative efficiency (RE) of 88.73% in residual variance compared to the RCBD. This improvement in precision resulted in specific changes in the ranking of the family evaluated, demonstrating the method's ability to prevent the discarding of promising genotypes. In experiment 2, RCBD was efficient in ensuring residual independence, outperforming neighbor adjustments in precision. It was concluded that the use of neighbor methods, combined with the study of spatial structure, is a powerful strategy to mitigate low local control efficiency when is needed.

Keywords: breeding programs; spatial dependency; spatial covariance; experimental precision; moving averages; Papadakis

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b> - Configurações de vizinhança consideradas por Feres (2009) .....	12
<b>Figura 2</b> - Instituições Federais de Ensino Superior participantes da RIDESA.....	14
<b>Figura 3</b> - Parâmetros de um semivariograma experimental com patamar ( $C$ ), efeito pepita ( $C_0$ ) e alcance ( $a$ ) .....	17
<b>Figura 4</b> - Representação gráfica de semivariogramas experimentais e modelos teóricos com patamar ( $C$ ), efeito pepita ( $C_0$ ) e alcance ( $a$ ).....	18
<b>Figura 5</b> - Funções de covariância e semivariograma com patamar ( $C$ ), efeito pepita ( $C_0$ ) e alcance ( $a$ ) .....	20
<b>Figura 6</b> - Esquema ilustrativo da seleção de parcelas vizinhas a partir de uma parcela de referência com base no alcance efetivo .....	25

## LISTA DE TABELAS

<b>Tabela 1</b> - Experimento, número de blocos (b), número de famílias (F), número total de parcelas avaliadas (NP), espaçamento (E), número de sulcos por parcelas (S), comprimento das parcelas em metros (C).....	23
<b>Tabela 2</b> - Índices ambientais calculados pelos métodos de médias móveis e Papadakis ponderados por informações espaciais .....	25
<b>Tabela 3</b> - Análise de variância para avaliação de 60 famílias de cana-de-açúcar do experimento 1 em DBC .....	31
<b>Tabela 4</b> - Índices de Dependência Espacial (IDE) e Índice de Moran para os resíduos estimados via DBC e DIC do experimento 1, com classificação segundo Zimback (2010) .....	32
<b>Tabela 5</b> - Parâmetros de ajuste e critérios de seleção dos modelos espaciais para o experimento 1 em DBC .....	33
<b>Tabela 6</b> - Parâmetros de ajuste e critérios de seleção dos modelos espaciais para o experimento 1 em DIC .....	34
<b>Tabela 7</b> - Valores do Índice de Moran para os resíduos sob diferentes métodos de ajuste nos delineamentos DBC e DIC .....	35
<b>Tabela 8</b> - Comparação da qualidade experimental entre os modelos estatísticos tradicionais e espaciais no experimento 1 .....	37
<b>Tabela 9</b> - Ranqueamento das 10, 20 e 30% das melhores famílias de cana-de-açúcar do experimento 1, com base nas médias de TCH estimadas para cada o modelo de análise do DBC e PPC .....	39
<b>Tabela 10</b> - Análise de variância para avaliação de 98 famílias de cana-de-açúcar do experimento 2 na área experimental da RIDESA em DBC .....	40
<b>Tabela 11</b> - Índices de Dependência Espacial (IDE) e índice de Moran para os resíduos estimados na análise em DBC e DIC do experimento 2, com classificação segundo Biondi et al. (1994).....	41
<b>Tabela 12</b> - Parâmetros de ajuste e critérios de seleção dos modelos espaciais para o experimento 2 em DIC .....	42
<b>Tabela 13</b> - Valores do Índice de Moran para diferentes tipos de análises aplicadas ao experimento 2.....	43
<b>Tabela 14</b> - Comparação da qualidade experimental entre os modelos tradicionais e espaciais no experimento 2.....	45
<b>Tabela 15</b> - Ranqueamento das 10, 20 e 30% das melhores famílias de cana de açúcar do experimento 2, com base nas médias de TCH estimadas para cada o modelo de análise do DBC e PPID .....	46

## SUMÁRIO

1. INTRODUÇÃO .....	10
2. REVISÃO DE LITERATURA .....	14
2.1. Programa de Melhoramento Genético da cana-de-açúcar.....	14
2.2. Delineamento em Blocos Casualizados .....	15
2.3. Dependência espacial .....	15
2.4. Métodos de Vizinhança .....	20
2.4.1. Médias móveis e Papadakis.....	21
3. MATERIAIS E MÉTODOS.....	23
3.1 Dados experimentais.....	23
3.2 Modelos.....	24
3.3 Análises estatísticas.....	27
3.4 Avaliação das metodologias.....	29
4. RESULTADOS E DISCUSSÃO.....	31
4.1. Análise do experimento 1 .....	31
4.2 Análise do experimento 2 .....	40
5. CONCLUSÕES .....	47
REFERÊNCIAS.....	48

## 1. INTRODUÇÃO

Típica de climas tropicais e subtropicais, a cana-de-açúcar (*Saccharum spp.*) é uma gramínea semi-perene originária do sudeste asiático. Como a maior parte do território brasileiro encontra-se entre o Equador e o Trópico de Capricórnio, suas condições de clima e solo foram favoráveis ao plantio da cultura. Sua introdução no país ocorreu em 1532, inicialmente no litoral e, posteriormente, nas áreas interioranas com a expansão da colonização portuguesa (CARMO, 2023; RODRIGUES e ROSS, 2020).

A princípio, o cultivo da cana-de-açúcar era somente para a extração da sacarose. Nos engenhos, a planta era moída e o extrato resultante era fervido até criar consistência suficiente para ser cristalizado e exportado. No entanto, com a crise do petróleo na década de 1970, a procura por combustíveis alternativos tornou-se estratégica para países em desenvolvimento. No Brasil não foi diferente, uma vez que desde 1950 sua principal malha de transporte passou a ser a rodoviária. O desenvolvimento de biocombustíveis era, sobretudo, essencial para o pleno crescimento do país. A resposta à crise veio com a criação do Programa Nacional de Álcool (PROALCOOL), cuja finalidade era substituir os veículos movidos a gasolina por veículos movidos a álcool (CARMO, 2023; CARVALHO *et al.*, 2013).

Desde então, a cana-de-açúcar ganhou protagonismo no comércio global para a produção de etanol e geração de energia proveniente do bagaço. O Brasil desempenha papel fundamental nesse cenário por ser o maior produtor mundial da cultura (EMBRAPA, 2023). Vale ressaltar que essa liderança não é produto exclusivo de suas condições edafoclimáticas favoráveis. Os programas de melhoramento genético são determinantes para esse sucesso, ao proporcionarem lançamento de cultivares de cana-de-açúcar mais adaptadas às condições climáticas do país e resistentes a pragas, doenças e estresse hídrico.

Camargo (1998) comenta que, atualmente, grande parte dos híbridos de cana-de-açúcar utilizados comercialmente é proveniente de cruzamentos entre as espécies *Saccharum spontaneum* e *Saccharum officinarum*. Esses cruzamentos geraram plantas com elevada resistência a pragas e doenças, além do aumento no acúmulo de sacarose. Concomitantemente, a redução do teor de fibra torna os colmos mais macios, o que facilita a extração do caldo nas usinas.

O processo de lançamento de uma cultivar no mercado, no entanto, é complexo e demanda um longo período até a sua conclusão, que pode variar de 10 a 15 anos. Nesse intervalo, etapas como a determinação de genitores, realização de cruzamentos, processo de seleção de genótipos e avaliação em campo são indispensáveis para o sucesso do programa de melhoramento (COSTA; BUENO FILHO; RAMALHO, 2005). Trata-se, portanto, de um processo de alto custo, pois demanda investimentos a longo prazo em infraestrutura, tecnologia e recursos humanos. Portanto, desde o planejamento dos ensaios, é necessário adotar medidas para minimizar o efeito do erro experimental, conforme os princípios básicos da experimentação (repetição, casualização e controle local, caso necessário), propostos por Ronald A. Fisher (FISHER, 1966).

O delineamento em blocos casualizados (DBC) apresenta os três princípios básicos e aplica o controle local por meio da formação de blocos nos quais os tratamentos aparecem uma única vez. Espera-se que possua homogeneidade dentro de cada bloco e heterogeneidade entre os blocos, que pode ser pequena ou grande (GOMES, 1985). No entanto, ao se trabalhar com o DBC no melhoramento da cana-de-açúcar ele pode apresentar falhas que resultem na diminuição da precisão experimental, que deve ser evitada a todo custo em um programa de melhoramento. Essas falhas são provenientes do pouco material de genótipos disponível, que resulta em parcelas pequenas e em um número reduzido de repetições dos tratamentos (FASOULA; IOANNIDES; OMIROU, 2020). Além disso, a alta quantidade de progênies sendo avaliadas acarretam na necessidade de utilizar blocos extensos em campo, o que pode comprometer a uniformidade dentro do bloco (COSTA; BUENO FILHO; RAMALHO, 2005; FERES, 2009).

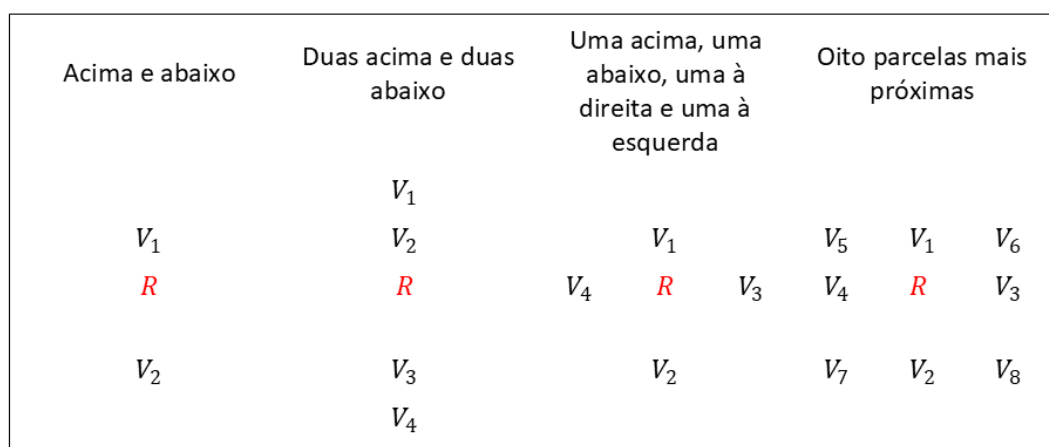
Tendo em vista esses desafios, Costa, Bueno Filho e Ramalho (2005) sugerem como alternativa ao melhorista o uso de análises espaciais, que consideram a posição de cada parcela no campo experimental, além de métodos de vizinhança, como médias móveis e o procedimento de Papadakis. Tais abordagens permitem ajustar os efeitos de parcelas vizinhas e capturar a estrutura espacial existente. Em seu estudo analisando a precisão experimental de ensaio de trigo, Storck e Silva (2014) concluíram que o método de Papadakis apresentou eficiência superior quando comparado ao DBC, delineamento no qual as cultivares foram plantadas. Cargnelutti

Filho *et al.* (2003), Storck *et al.* (2008) e Storck *et al.* (2009), também constataram a contribuição da análise de vizinhança na redução da variância espacial.

Embora a utilização de médias móveis e do método de Papadakis como covariáveis sejam alternativas para a redução da variância espacial, não existe consenso sobre a aplicação dessas metodologias. Na literatura, observam-se diferenças tanto na quantidade de vizinhos selecionados e na direção de sua escolha, quanto na forma de trabalhar com parcelas localizadas nas bordas do campo experimental, isto é, aquelas que não possuem o número completo de vizinhos definido. Candido *et al.* (2009) ressaltam que o número de vizinhos a ser considerado na análise é questionável, podendo variar de local para local e de experimento para experimento.

Feres (2009) em seu trabalho avaliou diferentes maneiras de definir as parcelas vizinhas de uma parcela de referência (R), conforme ilustrado na Figura 1. Já Costa, Bueno Filho e Ramalho (2005), considerou apenas quatro parcelas vizinhas a parcela de referência, uma acima, uma abaixo, uma à direita e uma a esquerda.

**Figura 1** - Configurações de vizinhança consideradas por Feres (2009)



**Fonte:** Adaptado de Feres (2009).

Candido *et al.* (2009), consideram, em suas análises, para uma das covariáveis, como vizinhas, as duas parcelas laterais, uma à direita e outra à esquerda e, para a outra covariável, as quatro parcelas adjacentes, da mesma maneira que Costa, Bueno Filho e Ramalho (2005). Para as parcelas que se localizavam nos extremos dos blocos, eles não consideraram a borda. Nesses casos, a média móvel e procedimento de Papadakis da parcela de referência foram calculadas, a depender de sua

localização, com apenas uma, duas ou três parcelas vizinhas, conforme a posição ocupada no campo.

Lucio *et al.* (2016), no estudo sobre tamanho de parcelas utilizadas em experimentos com a cultura da alface, adotaram, para o cálculo das covariáveis de Papadakis, as duas, quatro e oito parcelas mais próximas. Entretanto, ao calcularem o valor de Papadakis da parcela de referência, eles também incluíram seu resíduo. Para as parcelas nas extremidades, eles seguiram a mesma metodologia que Candido *et al.* (2009) seguiram.

Diante dessas limitações e diferenças metodológicas, o presente trabalho tem por objetivo avaliar uma abordagem para a aplicação de médias móveis e do método de Papadakis, visando melhorar a precisão experimental de ensaios no Delineamento em Blocos Casualizados (DBC). Essa estratégia fundamenta-se na incorporação de informações espaciais nos modelos, ao considerar a distância e a covariância espacial entre parcelas como ponderadores. Para tanto, foram utilizados dados de dois experimentos de famílias de cana-de-açúcar da RIDESA. O diferencial da metodologia consiste na utilização do semivariograma para identificar a estrutura de dependência espacial e definir, por meio do alcance efetivo, o número das parcelas vizinhas para o ajuste do modelo.

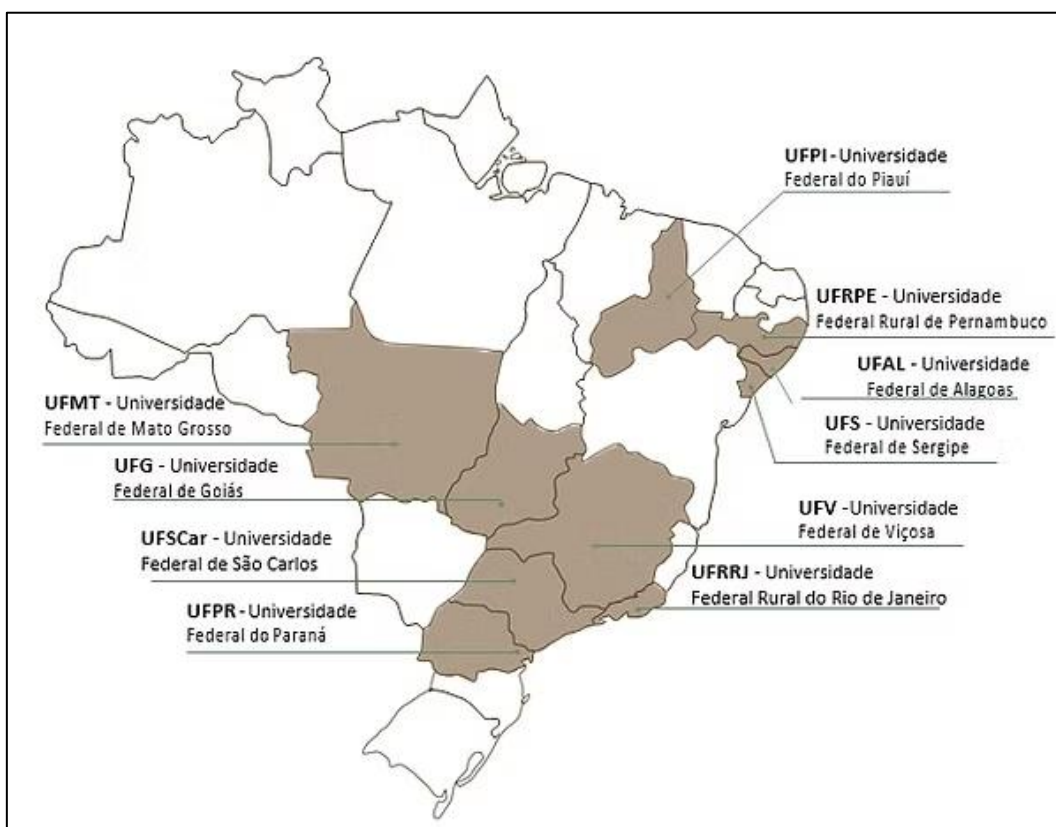
## 2. REVISÃO DE LITERATURA

### 2.1. Programa de Melhoramento Genético da cana-de-açúcar

Possuindo em torno de 10 milhões de hectare em área plantada de cana-de-açúcar, o Brasil é o maior produtor da cultura no mundo, sendo ela a segunda maior commodity do agronegócio brasileiro, ultrapassada apenas pela soja (FAO, 2020). Para além das condições de clima e solo, que favoreceram o bom desempenho da cultura, destacam-se os programas de melhoramento genético que desenvolvem cultivares mais produtivas, com maior tolerância ao estresse hídrico e resistência a pragas e doenças (CARMO, 2023).

No Brasil, o setor de inovações da cultura é predominantemente conduzido pelas instituições: Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (RIDESA), que lança no mercado às variedades RB Figura 2; o Instituto Agrônomo de Campinas, que desenvolve às variedades IAC; Centro de Tecnologia Canavieira, com as variedades CTC (MORAIS *et al.*, 2015).

**Figura 2** - Instituições Federais de Ensino Superior participantes da RIDESA



Fonte: RIDESA (2025). Disponível em <https://www.ridesa.com.br/>.

Diversos fatores podem influenciar no sucesso do melhoramento genético, como, por exemplo, a escolha adequada do delineamento experimental e a sua correta aplicação, a determinação das características a serem avaliadas e sua época de avaliação. As contribuições do melhoramento genético são diversas, com o desenvolvimento de variedades mais rústicas e com diferentes épocas de maturação, resistentes a pragas e doenças, além da ampliação do período de colheita e diversificação genética dos canaviais.

## **2.2. Delineamento em Blocos Casualizados**

No delineamento em blocos casualizados (DBC) os três princípios básicos da experimentação, repetição, casualização e controle local são utilizados. Os tratamentos são separados em blocos que contêm uma única vez cada um deles. A repetição possibilita a estimação do erro experimental ao aplicar o mesmo tratamento em diferentes unidades experimentais. A casualização assegura que as unidades experimentais recebam os tratamentos de forma aleatória, evitando tendências sistemáticas no experimento e garantindo a legitimidade da estimativa do erro experimental obtido pela repetição. O controle local, por sua vez, tem por objetivo diminuir a variância experimental por meio da formação de blocos homogêneos, dentro dos quais os tratamentos são comparados (GOMES, 1985; MONTGOMERY, 2017).

## **2.3. Dependência espacial**

Feres (2009) define a dependência espacial como sendo o grau de que o valor observado de uma variável com localização espacial definida (georreferenciada) tem de ser mais semelhante aos valores observados em sua vizinhança do que aos valores distantes. Atualmente, o semivariograma é a principal ferramenta utilizada para descrever essa dependência. Trata-se de uma função que associa distância com semivariância e que, ao quantificar a variabilidade dos dados em relação a sua distância de separação, indica até qual raio (alcance) as observações ainda se influenciam.

Para se determinar o valor da semivariância entre  $Z(X)$  e  $Z(X + h)$ , que estão separados por um vetor distância  $h$ , utiliza-se a equação (CAMARGO, 1998):

$$\gamma(h) = \frac{1}{2} E\{[Z(X + h) - Z(X)]^2\} \quad (1)$$

O estimador do semivariograma calcula a metade da média das diferenças quadráticas entre os valores de pares de observações da variável de interesse que estão separados por um vetor de distância  $h$ . Ele é determinado por (CAMARGO, 1998):

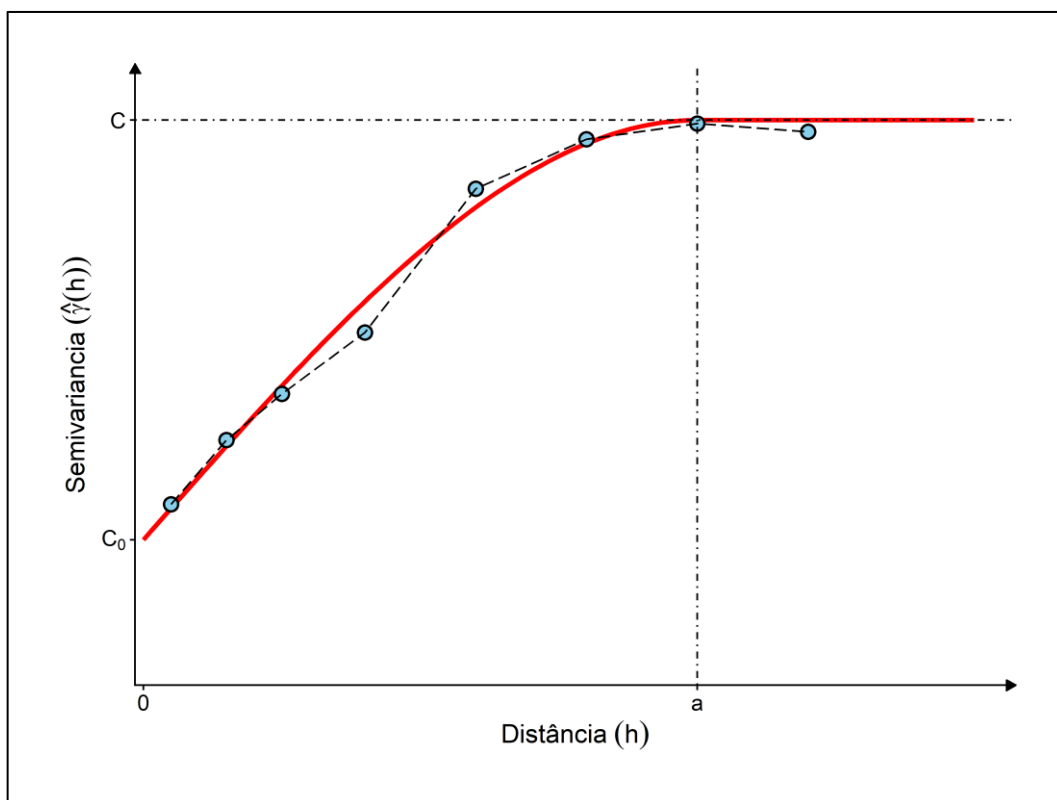
$$\hat{\gamma}(h) = \frac{1}{2 N(h)} \sum_{i=1}^{N(h)} \{[Z(x_i + h) - Z(x_i)]^2\} \quad (2)$$

com  $N(h)$  sendo o número de pares de valores medidos.

A figura 3 apresenta um semivariograma experimental com seus principais parâmetros: Patamar, Alcance e Efeito Pepita. O Alcance ( $a$ ) é a distância máxima dentro da qual as amostras ainda apresentam autocorrelação espacial. O Patamar ( $C$ ), por sua vez, é o valor máximo da semivariância a partir do qual se considera que não existe mais estrutura de dependência espacial. Por fim, o Efeito Pepita ( $C_0$ ) é o valor da semivariância para  $h = 0$  e representa a componente aleatória da variabilidade espacial (DRUCK *et al.*, 2004). Esse efeito pode vir de erros de medição ou de variações que ocorrem em distâncias menores do que o espaçamento mínimo escolhido (FERREIRA, 2008).

Conforme a Figura 3, à medida que a distância entre as observações cresce, a semivariância também aumenta até o momento em que o gráfico atinge o patamar (estabiliza) e, a partir dele, deixa de existir a dependência espacial entre as observações. Essa estabilização pode ou não ser assintoticamente atingida. Na figura 3, por exemplo, o patamar é assintoticamente atingido. Já no semivariograma do modelo esférico, representado em verde da figura 4, o patamar é alcançado em uma distância determinada pelo modelo matemático associado a ele.

**Figura 3** - Parâmetros de um semivariograma experimental com patamar ( $C$ ), efeito pepita ( $C_0$ ) e alcance ( $a$ )

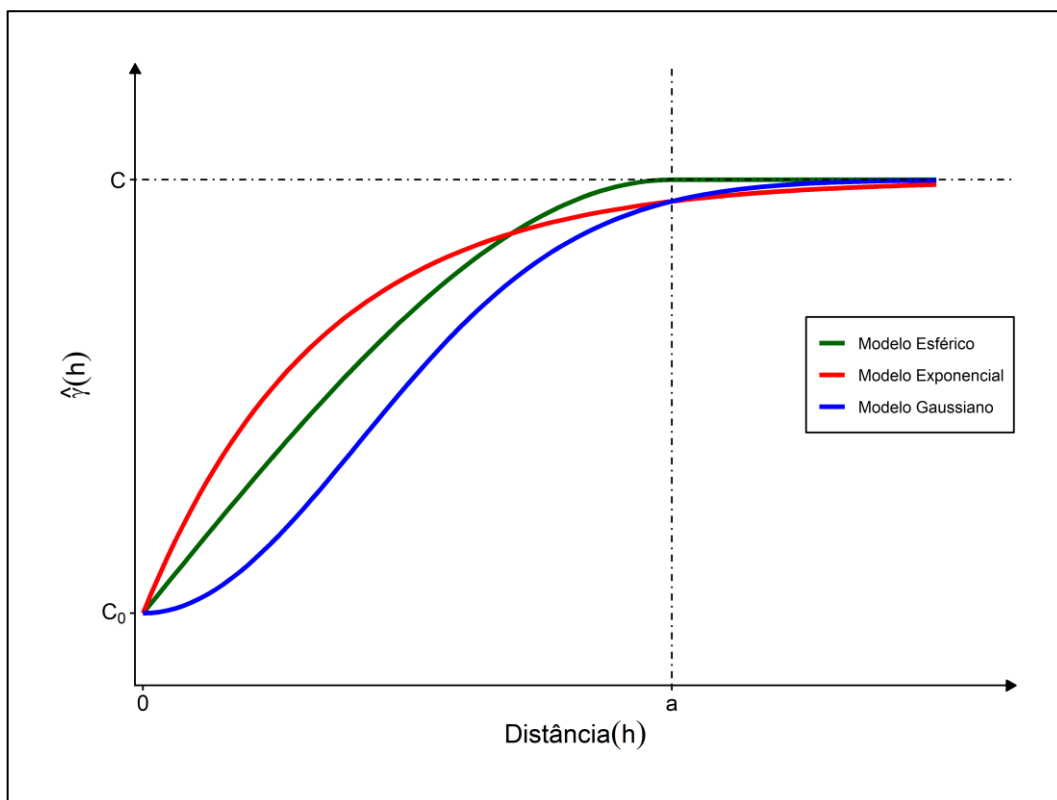


**Fonte:** Elaborada pelo autor.

Tendo obtido o semivariograma experimental, deve-se realizar o ajuste de algum modelo teórico para a curva, com base no padrão aparente dos pontos. Os métodos mais utilizados para o ajuste são: métodos da máxima verossimilhança (Maximum Likelihood - ML) e máxima verossimilhança restrita (Restricted Maximum Likelihood - REML); métodos de quadrados mínimos ordinários (Ordinary Least Squares - OLS) e método de quadrados mínimos ponderados (Weight Least Squares - WLS) (SOARES, 2000).

Os modelos teóricos que são frequentemente utilizados em análises espaciais são o esférico, exponencial e gaussiano. A representação gráfica desses modelos é apresentada na figura 4.

**Figura 4** - Representação gráfica de semivariogramas experimentais e modelos teóricos com patamar ( $C$ ), efeito pepita ( $C_0$ ) e alcance ( $a$ )



**Fonte:** Elaborada pelo autor.

Todos os três modelos são dados em função de 3 parâmetros, o patamar ( $C$ ), o efeito pepita ( $C_0$ ) e o alcance ( $a$ ), como evidenciado nas equações 3, 4 e 5 (CAMARGO, 1998; SOARES, 2000).

Modelo esférico (Esf)

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[ \frac{3}{2} \left( \frac{h}{a} \right) - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right], & \text{se } 0 < h < a \\ C_0 + C_1, & \text{se } h \geq a \end{cases} \quad (3)$$

Modelo exponencial (Exp)

$$\gamma(h) = C_0 + C_1 \left[ 1 - \exp \left( - \left( \frac{h}{a} \right) \right) \right], \quad (4)$$

Modelo Gaussiano (Gau)

$$\gamma(h) = C_0 + C_1 \left[ 1 - \exp \left( - \left( \frac{h}{a} \right)^2 \right) \right]. \quad (5)$$

Os modelos esférico e exponencial apresentam um crescimento rápido na origem, em contraponto ao modelo gaussiano, que tem um crescimento mais lento (figura 4). Assim, para fenômenos de natureza irregular (com maior variabilidade local), os dois primeiros modelos tendem a se ajustar melhor aos dados. Já para fenômenos mais regulares e contínuos, modelos Gaussianos se adequam melhor aos dados (SOARES, 2000).

Assim como para o semivariograma teórico, por meio do padrão aparente dos pontos no semivariograma experimental, pode-se escolher, entre os modelos de covariância existentes, o que se adequa mais com a dispersão dos pontos. A função covariograma retorna o valor da covariância entre os pares de valores da variável em estudo separados por uma distância  $h$  (FERREIRA, 2008). Seu estimador é dado por:

$$\hat{C}(h) = \frac{1}{|N(h)|} \sum_{i=1}^{N(h)} [Z(x_i) - \bar{Z}][Z(x_j) - \bar{Z}] \quad (6)$$

Sendo:

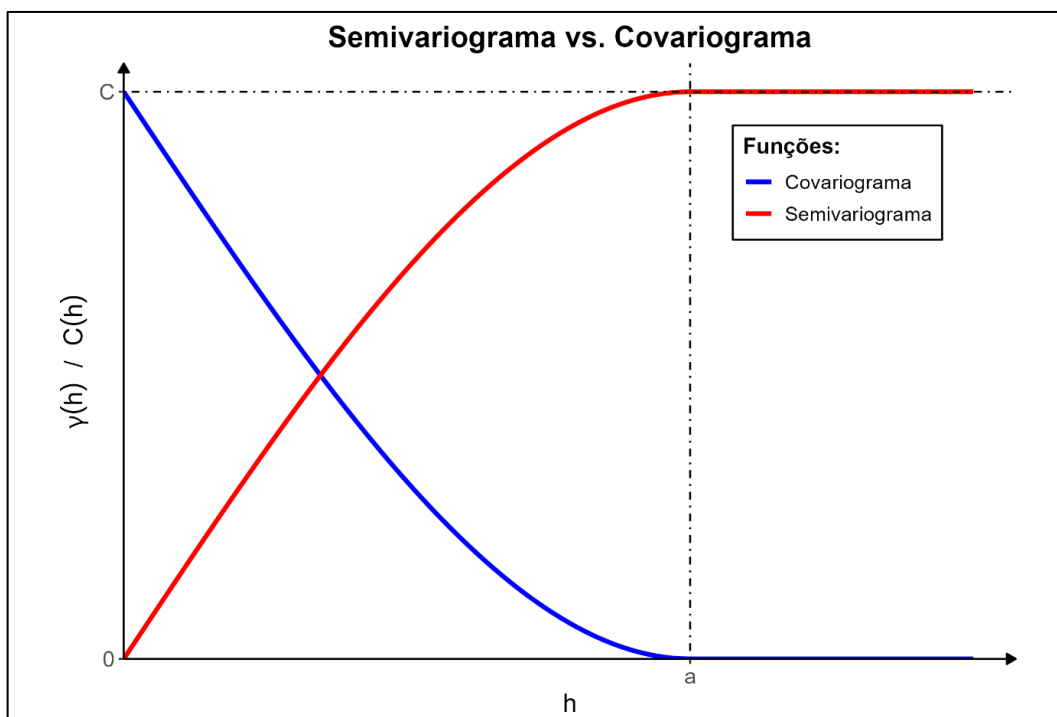
$$\bar{Z} = \frac{\sum_{i=1}^n Z(x_i)}{n} \quad (7)$$

O estimador da média.

Quando se considerada a estacionaridade de segunda ordem para os dados, a covariância e a semivariância se relacionam pela equação 8, sendo uma forma alternativa para caracterizar a dependência espacial entre os pares  $Z(x)$  e  $Z(x + h)$ . A figura 5 é permite visualizar a diferença entre o semivariograma e o covariograma.

$$\gamma(h) = C(0) - C(h) \quad (8)$$

**Figura 5** - Funções de covariância e semivariograma com patamar ( $C$ ), efeito pepita ( $C_0$ ) e alcance ( $a$ )



**Fonte:** Elaborada pelo autor.

## 2.4. Métodos de Vizinhaça

Os métodos de análise de vizinhaça procuram remover a dependência espacial entre as observações ou parcelas, por meio do uso de covariáveis obtidas ou pela média dos valores brutos das parcelas vizinhas ou pela média de seus resíduos (ZIMMERMAN; HARVILLE, 1991). Os métodos de vizinhaça tradicionais comumente utilizados na literatura são o de médias móveis e procedimento de Papadakis.

Souza *et al.* (2003), em seus estudos sobre alternativas experimentais em programas de melhoramento de eucalipto, elucidam que o resultado esperado na aplicação dessa metodologia é melhorar a precisão experimental ao eliminar parte das tendências existentes no solo por meio do desempenho das parcelas vizinhas. Em seu trabalho, os autores comparam a eficiência de diferentes delineamentos com os métodos de médias móveis e Papadakis, concluindo que ambos foram eficientes no controle da heterogeneidade ambiental intrabloco. Ao observar que tais metodologias

se equipararam ao látice nas análises de Souza *et al.* (2003), torna-se plausível supor que o ajuste por vizinhança possa atuar como um substituto eficaz para delineamentos com controle local.

#### 2.4.1. Médias móveis e Papadakis

No método de Papadakis (1937), o valor da variável observada em cada parcela é corrigido pelo efeito médio do tratamento aplicado a ela. Em seguida, é calculada para cada unidade experimental o seu respectivo índice ambiental, definido como o resíduo médio das parcelas vizinhas. Os índices obtidos são então utilizados como covariáveis, sendo incorporados ao modelo original por meio de uma análise de covariância.

Fórmula do índice ambiental de Papadakis ( $I_{PP_i}$ ):

$$I_{PP_i} = \frac{\sum_{j=1}^n \hat{e}_j}{n} \quad (9)$$

com  $n$  sendo o número de vizinhos.

Na metodologia de médias móveis (RICKEY, 1924), em contrapartida, o valor da variável observada em cada parcela é corrigido pela média dos valores das parcelas vizinhas. Nesse caso, o índice ambiental é definido como a média aritmética dos valores observados nas parcelas vizinhas a parcela de referência. Obtido o índice ambiental de cada parcela, ele é incorporado ao modelo como covariável e realiza-se uma análise de covariância.

Fórmula do índice ambiental Médias móveis ( $I_{MM_i}$ ):

$$I_{MM_i} = \frac{\sum_{j=1}^n y_j}{n} \quad (10)$$

com  $n$  sendo o número de vizinhos.

Ambas as metodologias podem ser aplicadas de forma complementar ao controle local convencional, conforme relatado nos estudos de Storck e Silva (2014) e Costa, Bueno Filho e Ramalho (2005). Essa integração permite que o ajuste de vizinhança capture a variância residual intrabloco, o que, em cenários com blocos extensos, condição frequente no melhoramento de cana-de-açúcar, possibilita a obtenção de estimativas de médias de tratamento mais precisas. Nos ensaios de Storck e Silva (2014), por exemplo, o número de cultivares variou entre 27 e 55, com parcelas úteis de 5  $m^2$ .

### 3. MATERIAIS E MÉTODOS

#### 3.1 Dados experimentais

Os dados utilizados nesse trabalho são provenientes de experimentos com famílias de cana-de-açúcar do Programa de Melhoramento Genético da Cana-de-Açúcar (PMGCA) da Universidade Federal de Viçosa em parceria com a RIDESA. Foram instalados dois experimentos no delineamento em blocos casualizados (DBC), conduzidos no Centro de Pesquisa e Melhoramento da Cana-de-Açúcar (CECA), localizado na Zona da Mata de Minas Gerais, município de Oratórios (494 m de altitude, latitude 20°25' S e longitude 42°48' W).

A Tabela 1 apresenta as características dos dois experimentos analisados: número de blocos, número de famílias, número total de parcelas, espaçamento em metros, número de sulcos por parcela e comprimento dos sulcos em metros.

**Tabela 1** - Experimento, número de blocos (b), número de famílias (F), número total de parcelas avaliadas (NP), espaçamento (E), número de sulcos por parcelas (S), comprimento das parcelas em metros (C)

Experimento	b	f	NP	E	Tamanho da parcela	
					S	C
1	4	60	240	1,4	2	4,5
2	4	98	392	1,4	2	1,5

**Fonte:** elaborada pelo autor.

Os genitores das famílias do experimento 1 foram, em sua maioria, das variedades RB, tendo apenas 7 com ou pai ou mãe da variedade CTC. O plantio ocorreu no dia 7 de março de 2018, tendo 10 plântulas por sulco, totalizando 20 delas por parcela. Não ocorreu perda de informações de parcelas e as análises foram realizadas com as informações das 240 parcelas.

Já no experimento 2, embora a maioria dos genitores fossem das variedades RB, também havia genitores das variedades CTC, UFPR, CV, SYN e SYNPLES. O plantio ocorreu no dia 10 de junho de 2021, tendo 12 plântulas por sulco, totalizando

24 delas por parcela. As análises foram realizadas com base nas informações de 388 parcelas, pois quatro parcelas foram perdidas.

As parcelas estavam regularmente espaçadas no campo experimental. Para a obtenção das coordenadas do centro de cada parcela, essenciais para a realização das análises posteriores, combinou-se as informações sobre o tamanho da parcela e do espaçamento entre elas. A distância entre parcelas foi determinada com base na distância euclidiana.

A variável utilizada para a avaliação dos genótipos foi a produtividade de colmos por hectare (TCH,  $Mg\ ha^{-1}$ ). A sua obtenção ocorreu por meio do produto entre a massa média e o número de colmos por metro; esse valor foi dividido pela área da parcela ( $m^2$ ) e, posteriormente, multiplicado pela constante 10 para a conversão da unidade para toneladas por hectare, conforme a metodologia descrita por Oliveira (2024).

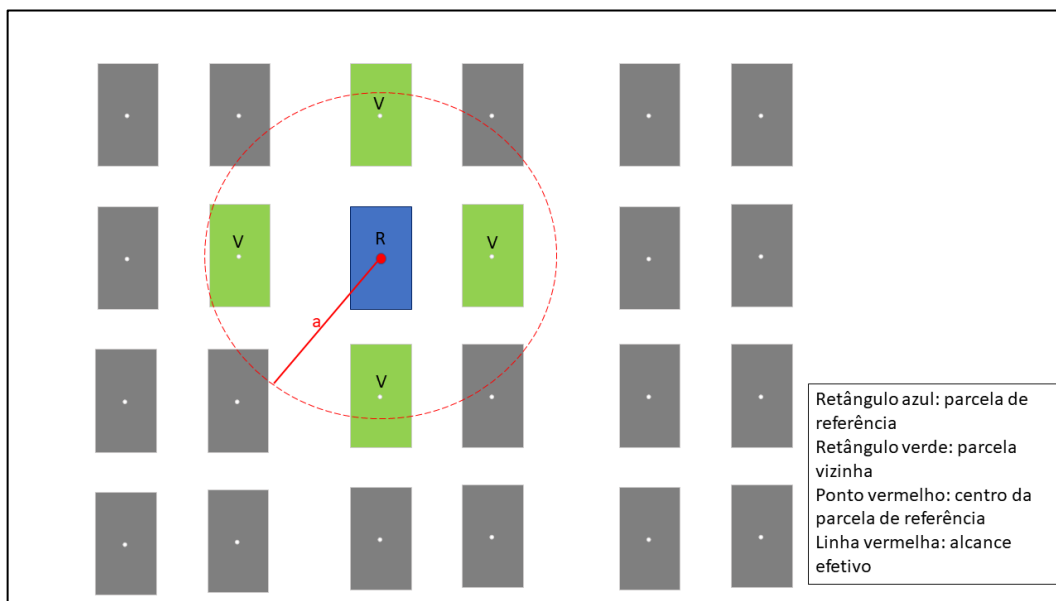
### 3.2 Modelos

Tanto o experimento 1 quanto o experimento 2 foram conduzidos sob o delineamento em blocos casualizados (DBC), sendo analisados considerando-se as estruturas de DBC e DIC. Para o ajuste espacial, foram aplicadas seis variações dos métodos de vizinhança, todas utilizando a estrutura de autocorrelação espacial do campo.

Em todas as metodologias, consideraram-se como vizinhas todas as parcelas situadas dentro do alcance efetivo, definido previamente pelo semivariograma. A distinção entre os métodos residu no critério de ponderação. Nos métodos médias móveis (MM) e Papadakis (PP), todas as parcelas dentro alcance contribuíram com o mesmo peso para o ajuste. Já nas demais variações, aplicaram-se ponderadores baseados no inverso da distância (MMID e PPID) e na covariância espacial (MMC e PPC).

A Figura 6 exemplifica esse critério de seleção de vizinhos baseado no raio de alcance  $a$ .

**Figura 6** - Esquema ilustrativo da seleção de parcelas vizinhas a partir de uma parcela de referência com base no alcance efetivo



Fonte: Elaborada pelo autor.

Os novos índices ambientais ponderados foram definidos com base em critérios geoestatísticos. Para a determinação das estimativas, considerou-se  $y_i$  como o valor observado na parcela  $i$  do campo experimental e  $\hat{e}_j$  o resíduo ambiental estimado da parcela vizinha  $j$ . Os termos  $Cov_{ij}$  e  $d_{ij}$  representam, respectivamente, a covariância espacial e a distância euclidiana de separação entre as parcelas  $i$  e  $j$ , enquanto  $n$  denota o número total de parcelas dentro do alcance estimado. As formulas para a obtenção dos índices ambientais ponderados (MMID, MMC, PPID e PPC) são descritas na tabela 2.

**Tabela 2** - Índices ambientais calculados pelos métodos de médias móveis e Papadakis ponderados por informações espaciais

<b>Índices Ambientais</b>			
$I_{MMID_i}$	$I_{MMC_i}$	$I_{PPID_i}$	$I_{PPC_i}$
$\frac{\sum_{j=1, j \neq i}^{n-1} \left( \frac{y_i}{d_{ij}} \right)}{\sum_{j=1, j \neq i}^{n-1} \left( \frac{1}{d_{ij}} \right)}$	$\frac{\sum_{j=1, j \neq i}^{n-1} Cov_{ij} \cdot y_j}{\sum_{j=1, j \neq i}^{n-1} Cov_{ij}}$	$\frac{\sum_{j=1, j \neq i}^{n-1} \left( \frac{\hat{e}_j}{d_{ij}} \right)}{\sum_{j=1, j \neq i}^{n-1} \left( \frac{1}{d_{ij}} \right)}$	$\frac{\sum_{j=1, j \neq i}^{n-1} Cov_{ij} \cdot \hat{e}_j}{\sum_{j=1, j \neq i}^{n-1} Cov_{ij}}$

Fonte: Elaborada pelo autor.

A seguir são descritos os modelos utilizados nesse trabalho, assumindo efeito fixo para todos os fatores:

$$\text{Modelo 1 (análise em DBC): } y_{ij} = \mu + f_i + b_j + e_{ij}$$

em que:

$y_{ij}$  é o valor observado da família  $i$  na repetição  $j$ ;

$\mu$  é a média de todas as famílias;

$f_i$  é o efeito do tratamento  $i$ ;

$b_j$  é o efeito do bloco  $j$ ;

$e_{ij}$  é o erro associado a família  $y_{ij}$ .

$$\text{Modelo 2 (análise em DIC): } y_{ij} = \mu + f_i + e_{ij}$$

Nos modelos 1 e 2 acima, o termo  $e_{ij}$  representa o erro experimental residual tradicional. Já nos modelos com ajuste de vizinhança, modelos 3 a 8, o erro residual  $\varepsilon_{ij}$  assume uma natureza distinta, uma vez que parte da variação ambiental que estava contida no erro aleatório é capturada pelo índice ambiental.

Os modelos 3 a 8 foram descritos com base no delineamento inteiramente casualizado. Entretanto, nos casos em que eles foram avaliados com base no DBC, acrescentou-se o efeito de bloco.

Em todos os modelos de 3 a 8, o parâmetro que  $\beta$  representa o coeficiente de regressão linear associado ao índice de vizinhança ( $I$ ).

$$\text{Modelo 3 (MM): } y_{ij} = \mu + f_i + \beta I_{MMij} + \varepsilon_{ij}$$

$$\text{Modelo 4 (MMID): } y_{ij} = \mu + f_i + \beta I_{MMIDij} + \varepsilon_{ij}$$

$$\text{Modelo 5 (MMC): } y_{ij} = \mu + f_i + \beta I_{MMCij} + \varepsilon_{ij}$$

$$\text{Modelo 6 (PP): } y_{ij} = \mu + f_i + \beta I_{PPij} + \varepsilon_{ij}$$

$$\text{Modelo 7 (PPID): } y_{ij} = \mu + f_i + \beta I_{PPIDij} + \varepsilon_{ij}$$

$$\text{Modelo 8 (PPC): } y_{ij} = \mu + f_i + \beta I_{PPCij} + \varepsilon_{ij}$$

### 3.3 Análises estatísticas

Todas as análises foram realizadas no software R (R Core Team, 2023), versão 4.3.2.

Obtidos os resíduos para os dois delineamentos (BDC e DIC), para se avaliar a normalidade e homocedasticidade dos resíduos foram aplicados, respectivamente, o teste de Shapiro – Wilk (SHAPIRO; WILK, 1965) e o teste de Bartlett (BARTLETT, 1937), que estão disponíveis nos pacotes base do R.

Para medir e avaliar a dependência espacial dos resíduos utilizou-se o Índice de Moran Global ( $I$ ), que mede a autocorrelação espacial de um conjunto de dados (BARBOSA *et al.*, 2013). Sua fórmula é dada por:

$$I = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

com,

$n$  sendo o número de observações

$x_i$  e  $x_j$  os valores da variável nas posições  $i$  e  $j$ .

$w_{ij}$  é o peso associado as observações  $u_i$  e  $u_j$

$S_0$  é o valor da soma de todos os pesos.

O Índice de Moran Global varia de -1 a 1. Valores próximos a zero indicam independência linear dos dados. Valores próximos a 1 indicam que valores semelhantes tendem a estar agrupados. Já os valores próximos a -1 indicam que valores opostos tendem a estar agrupados. Para avaliar a significância da autocorrelação espacial, é realizado um teste de hipóteses baseado na estatística  $Z$  (FERREIRA, 2015). A hipótese nula ( $H_0$ ) do teste presume que não há autocorrelação espacial dos valores. A hipótese alternativa ( $H_a$ ), por sua vez, presume que há autocorrelação espacial dos valores.

Para o cálculo do Índice de Moran Global, utilizou-se a metodologia dos K-vizinhos, considerando-se os 4 vizinhos mais próximos (SILVA, 2024). E seu teste de

significância foi realizado utilizando a função `moran.test()` do pacote `spdep` (BIVAND,2024).

Para os delineamentos em que o teste do Índice de Moran foram significativos, prosseguiu-se com a determinação do semivariograma experimental por meio da função `variog()`, do pacote `geoR` (RIBEIRO JR.; DIGGLE, 2024). Para o ajuste dos semivariogramas teóricos, optou-se pelos modelos Esférico, Exponencial e Gaussiano, que são os comumente utilizados na análise de dados da natureza dessa pesquisa. Utilizou-se a função `likfit()` para o ajuste dos parâmetros de cada modelo, que usa o método da máxima verossimilhança ou da máxima verossimilhança restrita.

Em relação a escolha do melhor modelo ajustado, utilizou-se o teste da Razão de verossimilhanças (LRT) (DUARTE, 2000), o Critério de Informação de Akaike (AIC) (AKAIKE, 1973) e a Raiz do Erro Quadrático Médio (RMSE) (HODSON, 2022).

Para os casos em que um dos três modelos (Esférico, Exponencial e Gaussiano) foi selecionado, determinou-se a covariância espacial por meio da criação de uma função que tinha como variáveis os parâmetros estimados do modelo escolhido.

O Índice de Dependência Espacial (IDE), apresentada em Biondi *et al.* (1994), foi calculado para identificar o grau de dependência espacial do modelo selecionado. Ele é dado pela expressão:

$$IDE = \frac{C}{C + C_o} \cdot 100\% \quad (12)$$

onde  $C_o$  é o efeito pepita e  $C$  a contribuição.

Considera-se, segundo Zimback (2010), os seguintes intervalos para o grau de dependência espacial:

- Forte dependência espacial caso  $IDE \geq 75\%$ ;
- Moderada dependência espacial caso  $25\% < IDE < 75\%$ ;
- Fraca dependência espacial se  $IDE \leq 25\%$ .

### 3.4 Avaliação das metodologias

Para a avaliação das metodologias apresentadas, utilizou-se a eficiência relativa (ER), o coeficiente de variação (CV%) e o Critério de Informação de Akaike (AIC).

A eficiência relativa (ER), é determinada pela razão entre o quadrado médio do resíduo da metodologia testada ( $QMR_{teste}$ ) e o quadrado médio da metodologia base ( $QMR_{base}$ ). Ela foi utilizada para comparar os métodos de análise de vizinhança propostos nesse trabalho com os modelos 1 e 2, conforme necessidade. A metodologia comparada é tanto melhor quanto menor for o valor obtido da ER.

A sua fórmula é dada por:

$$ER = \frac{QMR_{teste}}{QMR_{base}} \quad (13)$$

Utilizou-se o coeficiente de variação para comparar as metodologias analisadas quanto a precisão experimental. Ele é determinado pelo quociente entre o desvio-padrão residual e a média geral das observações, multiplicado por 100. Quanto menor o coeficiente de variação, mais precisa é a metodologia.

A sua fórmula é dada por:

$$CV\% = \frac{\sqrt{QMR}}{\bar{y}} \cdot 100 \quad (14)$$

A Acurácia Seletiva (ASE), também foi utilizada para avaliar a precisão experimental. Resende e Duarte (2007), comentam que ela não depende apenas do número de repetições e da variação residual, mas também do quociente entre as variações de natureza residual e genética associadas a variável resposta. Sua fórmula é dada por:

$$ASE = \left(1 - \frac{1}{F}\right)^{\frac{1}{2}} \quad (15)$$

Com F sendo a estatística para o efeito de família da análise de variância.

Os intervalos de precisão experimental para ASE utilizados neste trabalho foram os propostos no artigo de Campos *et al.* (2016):

- Moderada, caso  $0,60 \leq ASE < 0,80$ ;
- Alta, caso  $0,80 \leq ASE < 0,95$ ;
- Muito alta, caso  $ASE \geq 0,95$ .

### 3. 5 Ranqueamento das famílias

Para o ranqueamento das famílias, utilizou-se o pacote *emmeans* (Lenth, 2025) do R, que retorna os valores das médias corrigidas para cada família segundo o modelo utilizado e a função rank como feito por Ferreira (2020). Considerou-se 10%, 20% e 30% por cento das melhores famílias.

Para verificar a ordem de coincidência na classificação das famílias, foi utilizada a correlação de Spearman ( $r_s$ ), dada pela equação 15.

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)}, \quad (16)$$

com  $n$  sendo o número de observações e  $d_i$  a diferença entre os postos (ranks) de cada par de observações.

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Análise do experimento 1

**Tabela 3** - Análise de variância para avaliação de 60 famílias de cana-de-açúcar do experimento 1 em DBC

Fonte de Variação	gl	SQ	QMR	F calculado	p-valor
Família	59	119597	2027,1	3,68	<0,001*
Bloco	3	1404	467,97	0,85	0,468
Resíduo	177	97394	550,25	-	-
Total	239	218395	-	-	-

**Fonte:** elaborada pelo autor.

**Nota:** \* Significativo a 1% pelo teste F.

A tabela 3 apresenta o quadro da análise de variância (ANOVA), para o experimento 1 sob o delineamento em blocos casualizados (DBC). O efeito da família foi significativo, com p-valor < 0,001, o que evidencia diferenciações genéticas no material de estudo e permite, por conseguinte, a identificação de famílias superiores. No entanto, o efeito de bloco para a variável TCH não foi significativo, com p-valor = 0,468. Tal fato evidencia que a variação do solo ocorre de maneira diferente da planejada no delineamento, seja por uma possível homogeneidade da área ou mal posicionamento dos blocos. Em ambos os casos, a redução da eficiência do delineamento implica na perda de graus de liberdade do resíduo, fenômeno que, segundo Storck *et al.* (2008), compromete as estatísticas de precisão experimental.

Com relação à dependência espacial dos resíduos, os valores do Índice de Moran para os resíduos estimados pelo DBC e pelo DIC foram, respectivamente, 0,1739 e 0,1843, apresentando dependência espacial moderada ao nível de 1% de significância (Tabela 4). Esses resultados revelam que a divisão do campo experimental em quatro blocos foi, na verdade, insuficiente para captar a heterogeneidade ambiental da área. Devido a existência da autocorrelação espacial

nos dois delineamentos, procedeu-se com uma estratégia de análise de duas frentes: a aplicação do método de vizinhança sobre o DBC, para verificar possíveis ganhos de precisão, e sobre o DIC, para verificar a substituição do controle local.

**Tabela 4** - Índices de Dependência Espacial (IDE) e Índice de Moran para os resíduos estimados via DBC e DIC do experimento 1, com classificação segundo Zimback (2010)

Delineamento	Índice de Moran	IDE	Dependência Espacial
DBC	0,1739*	27%	Moderada
DIC	0,1843*	28%	Moderada

**Fonte:** elaborada pelo autor.

**Nota:** \* Significativo a 1% pelo teste Z.

Para modelar a estrutura de dependência espacial, foram ajustados modelos geoestatísticos para os resíduos das análises em DBC e em DIC, apresentados nas tabelas 5 e 6, respectivamente. Estas tabelas, apresentam as estimativas dos parâmetros para os modelos exponencial, esférico e gaussiano. Adicionalmente, constam os valores do AIC e o RMSE, utilizados como critério de seleção de melhor modelo, e o valor do Teste da Razão de Verossimilhança (LRT), que compara o modelo espacial (modelo com maior número de parâmetros ou modelo completo) com o modelo sem componentes espaciais (versão restrita), nesse caso, modelo de efeito pepita puro, que caracteriza a independência espacial entre os resíduos.

No que se refere aos resíduos estimados pelo DBC (Tabela 5), os modelos obtiveram valores semelhantes para AIC e RMSE. Entretanto, o Teste da Razão de Verossimilhança (LRT), apresentou significância apenas para os modelos exponencial e gaussiano, revelando superioridade em relação ao modelo de efeito pepita puro na explicação da estrutura de continuidade espacial dos dados. Entre os dois, o modelo gaussiano apresentou melhor ajuste, uma vez que obteve menores valores de AIC e RMSE. Essa escolha sugere a existência de uma continuidade suave da dependência espacial entre as parcelas ao longo do terreno (SOARES, 2000). Desse modo, constatou-se que parcelas que se encontravam dentro do raio de 9,68 m (alcance efetivo) no campo apresentavam dependência espacial, classificada como moderada segundo Zimback (2010). Feres (2009) observou valores de alcance efetivo

semelhantes para análises de progênies de feijão em que o índice de Moran indicou dependência espacial moderada.

**Tabela 5** - Parâmetros de ajuste e critérios de seleção dos modelos espaciais para o experimento 1 em DBC

Modelo	C <sub>1</sub>	a	C <sub>0</sub>	AIC	RMSE	LRT	p-valor
Efeito Pepita Puro	-	-	405,8	2129	20,23	-	-
Exponencial	180,60	10,62	227,9	2119	19,34	11,64	0,003*
Esférico	86,89	58,12	359,1	2124	19,61	6,98	0,03
Gaussiano	111,50	9,68	289,1	2118	19,24	12,81	0,001*

**Fonte:** elaborada pelo autor. Nota: C<sub>0</sub>: efeito pepita; C<sub>1</sub>: contribuição; a: alcance efetivo.

**Nota:** \* Significativo a 1% pelo LRT com 2 graus de liberdade.

De forma similar, no que se refere aos resíduos estimados pelo DIC (tabela 6), o modelo gaussiano foi o que apresentou melhor ajuste, com alcance efetivo de 9,91 m. O fato de o semivariograma ajustado ser o mesmo tanto para o DBC quanto para o DIC e os valores de seus parâmetros serem aproximados revela uma estrutura de dependência espacial análoga nos dois delineamentos, reiterando que a inclusão dos blocos foi insuficiente para filtrar a variação real do solo no experimento. Esse cenário justifica o uso de metodologias capazes de captar tal variação, como o método de Papadakis e médias móveis.

Vale ressaltar que os alcances efetivos de 9,68m e 9,91m, obtidos respectivamente para o DBC e DIC, englobam um número de parcelas superior ao considerado no cálculo do Índice de Moran, que utilizou as quatro unidades mais próximas. No campo experimental, esses alcances abrangeram, a depender da localização da parcela de referência, um máximo de 16 parcelas no DBC e 20 parcelas no DIC. Essa assimetria indica que a estrutura de autocorrelação espacial no campo possui abrangência superior à detectada por testes baseados em vizinhanças restritas, como observado em diversos estudos que utilizaram um número fixo de parcelas para correção da dependência espacial ( LUCIO *et al.*, 2016; STORCK *et al.*,

2008; STORCK; RIBEIRO; CARGNELUTTI FILHO, 2011; STORCK; SILVA, 2014). Tal observação fortalece a viabilidade de utilizar o alcance efetivo como critério para a delimitação de parcelas vizinhas em métodos de ajuste local, em particular os métodos de Papadakis e médias móveis, assegurando que toda a estrutura de autocorrelação espacial relevante seja considerada na análise.

**Tabela 6** - Parâmetros de ajuste e critérios de seleção dos modelos espaciais para o experimento 1 em DIC

Modelo	C <sub>1</sub>	a	C <sub>0</sub>	AIC	RMSE	LRT	p-valor
Efeito Pepita Puro	-	-	411,7	2132	20,37	-	-
Exponencial	178,3	11,92	237,0	2120	19,36	13,49	0,001*
Esférico	105,1	58,83	355,1	2124	19,58	10,12	0,006*
Gaussiano	118,1	9,91	297,8	2120	19,26	14,51	<0,001*

**Fonte:** elaborada pelo autor.

**Nota:** C<sub>0</sub>: efeito pepita; C<sub>1</sub>: contribuição; a: alcance efetivo. \* Significativo a 1% pelo LRT com 2 graus de liberdade.

Os valores do Índice de Moran revelaram que todas as metodologias de ajuste por vizinhança foram eficientes em assegurar a independência espacial dos resíduos nos dois delineamentos avaliados (Tabela 7). Observou-se um padrão de desempenho análogo entre o DBC e o DIC, onde os métodos de análise de vizinhança que utilizaram o inverso da distância e a covariância espacial como ponderadores apresentaram resultados superiores. Entre as variações do método de médias móveis, a que mais se destacou foi a MMC, com valores do índice de Moran iguais a 0,0127 e 0,0143 no DBC e DIC, respectivamente. Dos métodos de Papadakis, o PPID obteve melhor desempenho em ambos os cenários com valores do índice de Moran de -0,0305 e -0,0063.

**Tabela 7** - Valores do Índice de Moran para os resíduos sob diferentes métodos de ajuste nos delineamentos DBC e DIC

Método de análise	Índice de Moran	p-valor
Delineamento em Blocos Casualizados		
Original (DBC)	0,1739	< 0,001*
MM	0,0687	0,0588
MMID	0,0273	0,2497
MMC	0,0127	0,3583
PP	0,0471	0,1354
PPID	-0,0305	0,7142
PPC	-0,0683	0,9155
Delineamento Inteiramente Casualizado		
Original (DIC)	0,1843	< 0,001*
MM	0,086	0,0263
MMID	0,0364	0,1916
MMC	0,0143	0,3453
PP	0,0801	0,0350
PPID	-0,0063	0,5184
PPC	-0,0589	0,8799

**Fonte:** elaborada pelo autor.

**Nota:** \*Significativo a 1% pelo teste Z.

Esses resultados sugerem que tais métodos representam alternativas viáveis para a correção de falhas de blocagem ou, eventualmente, viabilizar a adoção do delineamento inteiramente casualizado (DIC) com correções a *posteriori*, como discutido por Storck *et al.* (2008). Ferreira (2020), ao avaliar propriedades físico-químicas do solo em experimentos com cana-de-açúcar, por meio de mapas de Krigagem destas propriedades, concluiu que a disposição dos blocos no campo não era condizente com a variabilidade real da área.

A adoção do delineamento inteiramente casualizado (DIC) com correções a *posteriori* por meio do ajuste espacial pode ser particularmente relevante em campos onde a variação de fertilidade é desconhecida ou em situações onde a utilização do DBC resultaria em blocos excessivamente grandes, cenário rotineiro em ensaios de seleção de genótipos, famílias e linhagens promissoras, que são caracterizados por um número elevado de genótipos e disponibilidade reduzida de material de propagação (FASOULA; IOANNIDES; OMIROU, 2020). Todavia, a substituição do controle local físico deve ser cautelosa. Como demonstrado posteriormente, nos resultados do Experimento 2, quando o delineamento em blocos é instalado corretamente, ele captura de forma eficiente a dependência espacial, tornando dispensáveis ajustes adicionais.

Em 100% das metodologias de vizinhança testadas ocorreram melhorias na precisão experimental nos dois delineamentos avaliados (Tabela 8). Dentre as metodologias de análise de vizinhança, aquelas que utilizaram o inverso da distância e a covariância espacial como ponderadores, apresentaram melhores estatísticas. Dentre elas, a metodologia PPC, sob ambos delineamentos, foi a mais eficiente, apresentando menor coeficiente de variação (CV), maior acurácia seletiva (ASE) e menor eficiência relativa (ER). A superioridade das modificações do método de Papadakis em relação às de médias móveis é esperada devido a heterogeneidade dentro dos blocos, uma vez que elas utilizam diretamente os resíduos como covariáveis, evitando ruídos adicionais, conforme discutido por Costa, Bueno Filho e Ramalho (2005).

Ao se comparar o DBC com o método PPC sob o DIC, observou-se a superioridade da metodologia em análise, com redução do Coeficiente de Variação de 19,70% para 18,54% e aumento da acurácia seletiva (ASE) de 0,8536 para 0,8716, além da redução de 11,27% na variância residual, conforme evidenciado pela eficiência relativa (ER). Essa superioridade endossa a discussão de Storck *et al.* (2008) sobre a possibilidade de descarte do controle local físico em favor da utilização de correções a *posteriori* por meio de métodos de vizinhança. Por outro lado, os resultados também sinalizam uma abordagem mais conservadora: a manutenção do DBC com a aplicação de correções espaciais apenas quando necessário. A viabilidade dessa estratégia é corroborada pelos resultados obtidos com o método

PPC sob o DBC que também apresentou resultados promissores (Tabela 8): o CV diminuiu 1,07% e a ER apresentou uma redução de 10,59% na variação residual.

**Tabela 8** - Comparação da qualidade experimental entre os modelos estatísticos tradicionais e espaciais no experimento 1

Tipo de análise	CV (%)	ER (%)	$r_s$	ASE
Delineamento em Blocos Casualizados				
DBC	19,70	-	-	0,8536
MM	19,16	94,68	0,99	0,8620
MMID	19,05	93,59	0,98	0,8637
MMC	19,01	93,05	0,98	0,8645
PP	19,10	93,96	0,99	0,8631
PPID	18,80	91,05	0,98	0,8676
PPC	18,63	89,41	0,98	0,8702
Delineamento Inteiramente Casualizado				
DIC	19,67	-	-	0,8539
MM	19,19	0,9510	0,99	0,8615
MMID	19,01	0,9328	0,99	0,8644
MMC	18,91	0,9237	0,98	0,8660
PP	19,20	0,9528	0,99	0,8615
PPID	18,81	0,9146	0,99	0,8674
PPC	18,54	0,8873	0,98	0,8716

**Fonte:** elaborada pelo autor.

**Nota:** CV: coeficiente de variação; ER: eficiência relativa;  $r_s$ : correlação de Spearman; ASE: acurácia seletiva.

Ganhos de precisão semelhantes foram observados por Storck e Silva (2014), que, seguindo abordagem similar à do presente estudo, compararam o desempenho

do delineamento em blocos com métodos de vizinhança. Ao utilizarem o método de Papadakis, os autores obtiveram, tanto em blocos homogêneos quanto heterogêneos, redução média do coeficiente de variação de 7,5% para 5,1% e aumento médio da acurácia seletiva de 0,90 para 0,95 em 48 ensaios de trigo originalmente instalados em DBC.

A eficiência dos métodos de vizinhança também é observada em outras culturas de médio e grande porte, como o milho (CARGNELUTTI FILHO; STORCK; LÚCIO, 2003; STORCK *et al.*, 2010), o eucalipto (SOUZA *et al.*, 2003) e a própria cana-de-açúcar (CANDIDO *et al.*, 2009). Neste contexto, Storck *et al.* (2010), estudaram o desempenho do procedimento de Papadakis em ensaios de competição de milho. Os autores relataram eficiência significativa nas estatísticas que descrevem a precisão experimental. A acurácia seletiva, por exemplo, aumentou de 0,82 para 0,92. Além disso, relataram que ensaios com apenas três repetições, quando analisados por esse método, permitiram identificar a produtividade de grãos com precisão de 86,5%. Tais resultados sugerem que, para culturas de portes semelhantes, o uso de informações de parcelas vizinhas compensa limitações no número de repetições, outro problema rotineiro no melhoramento genético.

As estimativas do coeficiente de correlação de Spearman ( $r_s$ ) apresentaram magnitude elevada para todas as metodologias testadas em comparação ao DBC, oscilando entre 0,98 e 0,99. Isto indica que praticamente não se verificaram grandes alterações no ordenamento de famílias. Contudo, a análise comparativa entre o DBC e o método PPC (Tabela 9) revelou alterações pontuais no ranqueamento. No estrato das 10% superiores (6 melhores famílias), a família 470 não seria selecionada pela metodologia do DBC. Para as 20% superiores (12 melhores famílias), a família 539 não seria selecionada. Por fim, para as 30% superiores (18 melhores famílias), a família 512 não seria selecionada. Tais resultados indicam que a omissão dos efeitos da dependência espacial pode levar ao descarte de famílias promissoras nas primeiras etapas do melhoramento genético.

**Tabela 9** - Ranqueamento das 10, 20 e 30% das melhores famílias de cana-de-açúcar do experimento 1, com base nas médias de TCH estimadas para cada o modelo de análise do DBC e PPC

Ranking	DBC	PPC	Raking	DBC	PPC	Raking	DBC	PPC
1	499	499	7	470	219	13	539	384
2	289	289	8	348	424	14	514	514
3	285	473	9	424	466	15	343	293
4	473	497	10	384	439	15	293	343
5	219	285	11	466	348	17	234	<b>512</b>
6	497	<b>470</b>	12	439	<b>539</b>	18	475	475

**Fonte:** elaborada pelo autor.

De forma análoga ao observado para o delineamento em blocos, as estimativas do coeficiente de correlação de Spearman ( $r_s$ ) apresentaram elevada magnitude para todas as metodologias testadas tendo o DIC como referência, com valores acima de 0,98. Quanto ao ordenamento das famílias, verificou-se que o DIC não divergiu do DBC, enquanto o PPC promoveu apenas uma permutação entre as famílias 439 e 466. Uma vez que ambas as famílias estavam dentro do estrato das 20% famílias superiores, a aplicação do ajuste espacial pelo PPC no DIC apresentaria alteração na seleção se fosse considerado apenas as 10% melhores famílias. Tal observação, fragiliza a possibilidade de descarte do controle local físico com uso de correções a *posteriori* discutida anteriormente.

As alterações observadas no ranqueamento das famílias ao utilizar o método PPC são condizentes com as discussões sobre sensibilidade de modelos de ajuste espacial presentes na literatura. Ferreira (2020), ao utilizar informações geoestatísticas provenientes de atributos do solo para o ajuste de modelos, compreende a reclassificação dos genótipos como um indicativo da eficiência do método. Para ele, tais ajustes, quando fundamentados na dependência espacial,

permitem a identificação de materiais superiores que seriam erroneamente descartados em análises convencionais.

Nessa mesma linha, Candido *et al.* (2009), ao estudarem a análise de vizinhança em genótipos de cana-de-açúcar, verificaram que os métodos de médias móveis e Papadakis aumentaram a eficiência do controle local. Com relação ao ordenamento dos genótipos, as metodologias apresentaram coeficientes de correlação de Spearman superiores a 0,72, reforçando a viabilidade dessas abordagens na seleção de materiais genéticos.

## 4.2 Análise do experimento 2

A tabela 10 apresenta o quadro da ANOVA para o experimento 2, conduzido sob o delineamento em blocos casualizados (DBC). O efeito de família foi significativo, constatando a existência de variabilidade genética no material de estudo e viabilizando a identificação de genótipos superiores. A significância observada para o efeito de blocos corrobora a presença de heterogeneidade ambiental no campo experimental, que aparentemente foi devidamente controlada por meio do controle local.

**Tabela 10** - Análise de variância para avaliação de 98 famílias de cana-de-açúcar do experimento 2 na área experimental da RIDESA em DBC

Fonte de Variação	gl	SQ	QMR	F calculado	p-valor
Família	97	185823	1915,7	2,48	<0,001*
Bloco	3	39567	13189	17,11	<0,001*
Resíduo	287	221279	771,0	-	-
Total	387	446669	-	-	-

**Fonte:** elaborada pelo autor.

**Nota:** \* Significativo a 1% pelo teste F.

A eficiência do DBC em conter a dependência espacial é confirmada pelos resultados do Índice de Moran (Tabela 11). Para esses resíduos, o índice foi de 0,0614, com p-valor > 1%, indicando independência espacial. Esse resultado se diferencia do obtido no experimento 1, no qual o controle local foi incapaz de assegurar a independência espacial dos resíduos. Tal distinção reforça que, para o presente experimento, a divisão do campo experimental em quatro blocos foi suficiente para absorver as variações ambientais do terreno, corroborando as observações de Matei *et al.* (2017), que também relataram a eficiência do controle local físico em ensaios de soja.

Diante da eficiência dos blocos, optou-se por não realizar ajustes de vizinhança sobre o DBC, diferindo em parte da abordagem realizada no experimento 1. Neste ensaio, as metodologias espaciais foram aplicadas exclusivamente sobre os resíduos do DIC, objetivando verificar se tais métodos seriam capazes de igualar ou superar a eficiência do controle local.

**Tabela 11** - Índices de Dependência Espacial (IDE) e índice de Moran para os resíduos estimados na análise em DBC e DIC do experimento 2, com classificação segundo Biondi *et al.* (1994)

Delineamento	Índice de Moran	IDE	Dependência Espacial
DBC	0,0614	-	-
DIC	0,1823*	24%	Fraca

**Fonte:** elaborada pelo autor.

**Nota:** \* Significativo a 1% pelo teste Z.

Como os resíduos estimados pelo DIC apresentaram dependência espacial, mesmo que fraca, procedeu-se com o ajuste de modelos geoestatísticos para esses resíduos, dos quais os modelos exponencial e esférico ajustaram para todos os parâmetros (Tabela 13). O modelo gaussiano, entretanto, indicou ausência de estrutura espacial. Ao aplicar o teste da razão de verossimilhança (LRT), verificou-se que ele foi significativo para os modelos exponencial e esférico, que apresentaram valores de AIC e RMSE próximos. O modelo esférico foi selecionado por apresentar, ainda que minimamente, menores valores para esses critérios de seleção. Assim, observou-se que parcelas situadas dentro de um raio de 58,52 metros no campo experimental apresentavam dependência espacial.

**Tabela 12** - Parâmetros de ajuste e critérios de seleção dos modelos espaciais para o experimento 2 em DIC

Modelo	$C_1$	a	$C_0$	AIC	RMSE	LRT	p-valor
Efeito Pepita Puro	-	-	672,73	3633	26,00	-	-
Exponencial	192,60	67,95	529,00	3595	24,18	40,92	< 0,001*
Esférico	193,20	58,52	542,40	3593	24,15	42,62	< 0,001*

**Fonte:** elaborada pelo autor.

**Nota:**  $C_0$ : efeito pepita;  $C_1$ : contribuição; a: alcance efetivo. \* Significativo a 1% pelo teste da razão de verossimilhança com 2 graus de liberdade.

Os valores do Índice de Moran apresentados na Tabela 13 revelaram que, além do DBC, apenas as metodologias MMID e PPID foram eficientes em garantir a independência espacial dos resíduos, apresentando valores de índice ligeiramente inferiores à análise com controle local. Esse resultado sugere que, mesmo no caso em que os blocos eram heterogêneos, os métodos de vizinhança com a ponderação pelo inverso da distância foram tão eficientes quanto os blocos para assegurar a independência espacial dos resíduos, embora não tenham angariado ganhos de precisão experimental, como será discutido adiante.

**Tabela 13** - Valores do Índice de Moran para diferentes tipos de análises aplicadas ao experimento 2

Tipo de análise	Índice de Moran	p-valor
DBC	0,0614	0,032
DIC	0,1823	< 0,001*
MM	0,1778	< 0,001*
MMID	0,0498	0,071
MMC	0,1143	< 0,001*
PP	0,1870	< 0,001*
PPID	0,0067	0,397
PPC	0,0868	0,006*

**Fonte:** elaborada pelo autor.

**Nota:** \* Significativo a 1% pelo teste Z.

Todas metodologias de análise de vizinhança testadas foram menos eficientes que o DBC, apresentando maiores coeficientes de variação e menores valores de acurácia seletiva (Tabela 14). Em relação à eficiência relativa (ER), as metodologias que utilizaram o inverso da distância e a covariância espacial entre as observações como ponderadores obtiveram os melhores resultados entre as alternativas espaciais. Diferente do observado no experimento 1, a metodologia PPID obteve melhor desempenho, com coeficiente de variação igual 31,12%, alta acurácia seletiva (0,7622) e menor eficiência relativa (1,0402). Entre as médias móveis, a MMID destacou-se com a segunda menor eficiência relativa (1,0739) entre as demais metodologias.

A superioridade das metodologias baseadas no inverso da distância em relação às baseadas em covariância espacial abre margem para pelo menos duas

possibilidades que não necessariamente se excluem. A primeira indica que, em condições de dependência espacial fraca, modelos de vizinhança que utilizam a distância euclidiana entre parcelas como ponderador apresentam maior robustez. A segunda possibilidade sugere que a fraca dependência espacial, em conjunto com valores atípicos, pode ter comprometido o ajuste do semivariograma, resultando na superestimação de seus parâmetros.

Costa, Bueno Filho e Ramalho (2005), ao avaliarem métodos de vizinhança no melhoramento genético de feijão e milho em diversos ensaios, concluíram que, em média, o controle local via látice foi eficiente. Nos casos em que os métodos de análise de vizinhança (MM e PP) foram aplicados sem considerar o controle local, não houve ganho de eficiência em relação a análise tradicional, corroborando os resultados deste experimento. Por outro lado, Matei *et al.* (2017), avaliando 324 genótipos de soja, observaram que o método de Papadakis proporcionou melhorias na precisão experimental tanto em blocos homogêneos quanto heterogêneos, com acurácia seletiva superior a 90% em todos ensaios. Em concordância com Matei *et al.* (2017), Storck *et al.* (2008), avaliando também genótipos de soja, obtiveram ganhos de precisão, com diminuição média do coeficiente de variação de 3,64% e aumento médio de 18,2% na eficiência relativa, nos ensaios que apresentavam heterogeneidade entre blocos.

As estimativas do coeficiente de correlação de Spearman ( $r_s$ ) apresentaram magnitude elevada para todas as metodologias testadas em comparação ao DBC, variando de 0,96 a 0,99. Isto indica que, de modo geral, não ocorreram grandes alterações no ordenamento das famílias, assim como no primeiro ensaio. Contudo, a análise comparativa entre o DBC e o método PPID sob o DIC revelou alterações pontuais no ranqueamento (Tabela 15). Observou-se que para as 10%, 20% e 30% melhores famílias, ocorreria perda de famílias promissoras caso o PPID fosse deliberadamente adotado, resultando em prejuízos qualitativos ao programa de melhoramento.

**Tabela 14** - Comparação da qualidade experimental entre os modelos tradicionais e espaciais no experimento 2

Tipo de análise	CV (%)	ER (%)	$r_s$	ASE
DBC	30,51	-	-	0,7730
DIC	32,96	1,1674	0,99	0,7278
MM	32,97	1,1677	0,99	0,7277
MMID	31,62	1,0739	0,96	0,7532
MMC	32,05	1,1038	0,98	0,7454
PP	33,00	1,1698	0,99	0,7271
PPID	31,12	1,0402	0,98	0,7622
PPC	31,65	1,0765	0,99	0,7526

**Fonte:** elaborada pelo autor.

**Nota:** CV: coeficiente de variação; ER: eficiência relativa;  $r_s$ : correlação de Spearman; ASE: acurácia seletiva.

Esse resultado põe em xeque a viabilidade de se descartar o controle local confiando em ajustes posteriores por meio de métodos de vizinhança. Por outro lado, essas evidências ampliam a possibilidade de se utilizar a abordagem mais conservadora que consiste em adotar o delineamento em blocos casualizados e, em situações de baixa eficiência deles, avaliar a dependência espacial por meio do semivariograma. No caso de ter se detectado uma estrutura de dependência espacial nos resíduos, utilizar métodos de vizinhança a fim de melhorar a precisão experimental. Essa abordagem permite resgatar famílias superiores que poderiam ser subestimadas e descartadas por conta da falha do controle local, unindo o rigor do planejamento experimental à flexibilidade da modelagem espacial.

**Tabela 15** - Ranqueamento das 10, 20 e 30% das melhores famílias de cana de açúcar do experimento 2, com base nas médias de TCH estimadas para cada o modelo de análise do DBC e PPID

Raking	DBC	PPID	Raking	DBC	PPID	Raking	DBC	PPID
1	1485	1469	11	1268	1082	21	1282	1282
2	1469	1485	12	1472	1297	22	1277	1022
3	1465	1465	13	1297	1472	23	1022	<b>1031</b>
4	1484	1484	14	1455	1455	24	1228	1460
5	1456	1225	15	1262	<b>1228</b>	25	1460	1121
6	1225	1456	16	1082	<b>1475</b>	26	1307	<b>1271</b>
7	1264	<b>1262</b>	17	<b>1031</b>	1268	27	1015	1015
8	1272	1264	18	1235	1235	28	1121	1487
9	1457	1272	19	1232	<b>1277</b>	29	1487	<b>1110</b>
10	<b>1475</b>	1457	20	<b>1271</b>	1232	30	<b>1310</b>	1307

**Fonte:** elaborada pelo autor.

## 5. CONCLUSÕES

O experimento 1 apresentou dependência espacial moderada tanto sob o delineamento em blocos casualizados (DBC) quanto no delineamento inteiramente casualizado (DIC). Nesses cenários, os métodos de análise de vizinhança proporcionaram melhorias na precisão experimental, com valores próximos para as estatísticas de coeficiente de variação (CV), eficiência relativa (ER) e acurácia seletiva (ASE). Todas as metodologias obtiveram valores elevados de coeficiente de correlação de Spearman ( $r_s$ ), indicando que não ocorreram grandes diferenças na classificação geral entre as análises.

No experimento 2, sob o DBC, os blocos foram eficazes para garantir a independência dos resíduos. Por outro lado, sob o DIC, o ensaio apresentou fraca dependência espacial. Nesse contexto, os métodos de análise de vizinhança não superaram o DBC em termos de CV, ER e ASE. Assim como no experimento 1, as correlações de Spearman ( $r_s$ ) continuaram elevadas para todas as metodologias testadas.

Os resultados das análises dos dois ensaios indicam que, em cenários com baixa eficiência dos blocos, a utilização de métodos de vizinhança para melhorar a precisão experimental é uma estratégia eficiente, caso seja detectada estrutura de dependência espacial. Em contrapartida, em casos onde os blocos apresentam eficiência satisfatória, recomenda-se prosseguir com a análise via DBC. Portanto, a estratégia mais conservadora consiste em trabalhar, *a priori*, com o delineamento em blocos casualizados e realizar correções *a posteriori* por meio dos métodos de vizinhança, médias móveis e Papadakis, se necessário.

Dentre todas as metodologias de análise de vizinhança avaliadas, as que consideraram como ponderadores o inverso da distância e a covariância espacial apresentaram superioridade em relação às outras. São necessários novos experimentos para consolidar os resultados observados e avaliar o desempenho dessa metodologia de vizinhança proposta com a que utiliza um número fixo de vizinhos.

## REFERÊNCIAS

- ARRUDA, Hermano Vaz de. **Eficiência do delineamento em blocos ao acaso, em experiências comparativas de variedades e híbridos de milho**. *Bragantia*, v. 13, p. 217-222, 1954.
- BARBOSA, Isabelle Ribeiro et al. **Análise da distribuição espacial da tuberculose na região Nordeste do Brasil, 2005-2010**. *Epidemiologia e Serviços de Saúde*, v. 22, n. 4, p. 687-695, 2013.
- BARTLETT, Maurice S. **The statistical conception of mental factors**. *British journal of Psychology*, v. 28, n. 1, p. 97, 1937.
- BIONDI, F.; MYERS, D. E.; AVERY, C. C. **Geostatistically modeling stem size and increment in an old-growth forest**. *Canadian Journal of Forest Research*, v. 24, p. 1354-1368, 1994.
- BIVAND, R. S. **spdep: Spatial Dependence: Weighting Schemes, Statistics and Models**. *R package version 4.3.2*. 2024.
- CANDIDO, L. S.; PERECIN, D.; LANDELL, M. G. de A.; PAVAN, B. E. **Análise de vizinhança na avaliação de genótipos de cana-de-açúcar**. *Pesquisa Agropecuária Brasileira*, v. 4, p. 1304-1311, 2009.
- CAMARGO, Eduardo Celso Gerbi. **Geoestatística: fundamentos e aplicações**. In: **Geoprocessamento para projetos ambientais**. São José dos Campos: INPE, v. 29, 1998.
- CARGNELUTTI FILHO, Alberto; STORCK, Lindolfo; DAL'COL LÚCIO, Alessandro. **Ajustes de quadrado médio do erro em ensaios de competição de cultivares de milho pelo método de Papadakis**. *Pesquisa Agropecuária Brasileira*, v. 38, p. 467-473, 2003.
- CARMO, Andreza Sousa. **Propagação de genótipos de cana-de-açúcar por meio da técnica de mudas pré-brotadas**. 2023. Dissertação (Mestrado em Produção Vegetal) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Rio de Janeiro, 2023. Orientador: Silvio de Jesus Freitas.
- CAMPOS, Josmar Furtado de et al. **Classificação de famílias do feijoeiro sob diferentes cenários de dependência espacial e precisão experimental**. *Pesquisa Agropecuária Brasileira*, v. 51, p. 105-111, 2016.
- CARVALHO, José Ruy Porto de; SILVEIRA, Pedro Marques da; VIEIRA, Sidney Rosa. **Geoestatística na determinação da variabilidade espacial de características químicas do solo sob diferentes preparos**. *Pesquisa Agropecuária Brasileira*, v. 37, p. 1151-1159, 2002.
- CARVALHO, L. C.; BUENO, R. C. O. de F.; CARVALHO, M. M.; FAVORETO, A. L.; GODOY, A. F. **Cana-de-açúcar e álcool combustível: histórico, sustentabilidade e segurança energética**. *Enciclopédia Biosfera*, v. 9, n. 16, p. 530-542, 2013.

COSTA, Janaina Ribeiro; BUENO FILHO, Júlio Sílvio de Sousa; RAMALHO, Magno Antônio Patto. **Análise espacial e de vizinhança no melhoramento genético de plantas**. Pesquisa Agropecuária Brasileira, v. 40, p. 1073-1079, 2005.

DE OLIVEIRA, Ronaldo Pereira; GREGO, Célia Regina; BRANDÃO, Ziany Neiva. **Geoestatística aplicada na agricultura de precisão utilizando o Vesper**. Brasília: Embrapa, 2015.

DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. (ed.). **Análise espacial de dados geográficos**. Brasília: Embrapa, 2004.

EMBRAPA. CANA. Agência de informação tecnológica. Cultivos. 2023. Disponível em: <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/cana>.

FAO. FAOSTAT Statistical Database. **Food and Agriculture Organization of the United Nations**, 2020.

FASOULA, Dionysia A.; IOANNIDES, Ioannis M.; OMIROU, Michalis. Phenotyping and plant breeding: overcoming the barriers. **Frontiers in plant science**, v. 10, p. 1713, 2020.

FERES, A. L. G. **Análise Estatística na avaliação de produtividade no melhoramento genético do feijoeiro**. 2009. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Viçosa, Viçosa, MG.

FERREIRA, M. P. **Redução do adensamento amostral no ajuste de modelos de semivariogramas**. 2015. 52 f. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Viçosa, Viçosa, MG, 2015.

FERREIRA, M. P. **Geoestatística e aerofotogrametria aplicada à seleção de família de cana-de-açúcar**. 2020. 72 f. Tese (Doutorado em Estatística Aplicada e Biometria) - Universidade Federal de Viçosa, Viçosa, 2020.

FERREIRA, Sérgio Magalhães. **Análise de incertezas em modelos geométricos de depósitos minerais usando geoestatística e esquema de planejamento de experimentos**. 2008.

FISHER, Richard V. Rocks. **Composed of volcanic fragments and their classification**. Earth-Science Reviews, v. 1, n. 4, p. 287-298, 1966.

GALVÃO, João Antônio Martins. **Estimação da temperatura média na região amazônica por meio de técnicas de Cokrigagem**. 2012.

GOMES, Frederico Pimentel. **Curso de estatística experimental**. 11. ed. Piracicaba: Nobel, 1985.

LENTH, R. *emmeans: Estimated Marginal Means, aka Least-Squares Means*. Versão 1.11.2-8. **[Software de computador]**. [S. l.]: CRAN, 2025. DOI: 10.32614/CRAN.package.emmeans. Disponível em: <https://CRAN.R-project.org/package=emmeans>. Acesso em: [2024].

LUCIO, Alessandro D. et al. **Método de Papadakis e tamanho de parcela em experimentos com a cultura da alface**. Horticultura Brasileira, v. 34, p. 66-73, 2016.

MARCONATO, Renata; LAROCCA, Ana Paula C.; QUINTANILHA, José Alberto. **Análise do uso de tecnologias em estabelecimentos agropecuários por meio dos índices de Moran global e local**. Revista de Política Agrícola, v. 21, n. 1, p. 5-21, 2012.

MATEI, Gilvani et al. **Methods of analysis and number of replicates for trials with large numbers of soybean genotypes**. Ciência Rural, v. 47, n. 4, p. e20160629, 2017.

MORAIS, L. K. de; CURSI, D. E.; SANTOS, J. M. dos; SAMPAIO, M.; CAMARA, T. M. M.; SILVA, P. de A.; BARBOSA, G. V.; HOFFMANN, H. P.; CHAPOLA, R. G.; FERNANDES JUNIOR, A. R.; GAZAFFI, R. **Melhoramento Genético da Cana-de-Açúcar**. Aracaju: Embrapa Tabuleiros Costeiros, 2015. 40 p.

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 8. ed. New York: John Wiley & Sons, 2017.

PAPADAKIS, J. S. **Méthode statistique pour des expériences sur champ. Thésalonique: Institut d'Amélioration des Plantes à Salonique**, 1937. 30 p. (Bulletin, 23).

PLAZAS, Gloria M. R. et al. **Análise espacial da produtividade da cenoura irrigada em função das propriedades físicas do solo**. Revista Brasileira de Engenharia Agrícola e Ambiental, v. 28, p. e275518, 2024.

PONTES, José Marcelo; OLIVEIRA, Marcelo Silva de. **Uma proposta alternativa para a análise de experimentos de campo utilizando a geoestatística**. Ciência e Agrotecnologia, v. 28, p. 135-141, 2004.

SHAPIRO, Samuel Sanford; WILK, Martin B. An analysis of variance test for normality (complete samples). **Biometrika**, v. 52, n. 3-4, p. 591-611, 1965.

DE RESENDE, Marcos Deon Vilela; DUARTE, João Batista. **Precisão e controle de qualidade em experimentos de avaliação de cultivares**. Pesquisa Agropecuária Tropical, p. 182-194, 2007.

RIBEIRO JR., P. J.; DIGGLE, P. **geoR: Analysis of Geostatistical Data**. Versão 1.9-6. [Software de computador]. [S. l.]: CRAN, 2025. Disponível em: <https://CRAN.R-project.org/package=geoR>. Acesso em: [2024].

RICKEY, F. D. **Adjusting yields to their regression on moving average, as a means of correction for soil heterogeneity**. Journal of Agricultural Research, v. 27, p. 79-90, 1924.

RODRIGUES, G. S. S. C.; ROSS, J. L. S. **A trajetória da cana-de-açúcar no Brasil: perspectivas geográficas, histórica e ambiental**. Uberlândia: EDUFU, 2020. p. 253-269. ISBN: 978-65-86084-00-9.

SILVA, Mariana de Oliveira. **Análise da dependência espacial em experimentos com cana-de-açúcar da RIDESA**. 2024. 58 f. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Viçosa, Viçosa, MG, 2024.

SOARES, Amílcar. **Geoestatística para as ciências da terra e do ambiente**. Instituto Superior Técnico, 2000.

SOUZA, Elaine Aparecida de et al. **Alternativas experimentais na avaliação de progênies e clones em programas de melhoramento de eucalipto**. Revista Árvore, v. 27, p. 427-434, 2003.

STORCK, Lindolfo et al. **Utilização do método de Papadakis na melhoria da qualidade experimental de ensaios com soja**. Pesquisa Agropecuária Brasileira, v. 43, p. 581-587, 2008.

STORCK, Lindolfo et al. **Método de Papadakis e número de repetições em experimentos de soja**. Ciência Rural, v. 39, p. 977-982, 2009.

STORCK, Lindolfo; SILVA, Marcio Só. **Precisão experimental de ensaios de trigo analisados pelo método de Papadakis**. Ciência Rural, v. 44, p. 407-413, 2014.

ZIMBACK, C. R. L. **Análise espacial de atributos químicos de solo para fins de mapeamento da fertilidade do solo**. Botucatu, 2001. Tese (Livre-docência em Ciências Agrônômicas) - Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista, Botucatu, 2001.

ZIMMERMAN, Dale L.; HARVILLE, David A. **A random field approach to the analysis of field-plot experiments and other spatial experiments**. Biometrics, p. 223-239, 1991.