

## MODELAGEM ESPACIAL NA ANÁLISE DE UM PLANTIO EXPERIMENTAL DE CANDEIA

Cristina Henriques NOGUEIRA<sup>1</sup>  
Marcelo Silva de OLIVEIRA<sup>1</sup>  
Marcel Régis RAIMUNDO<sup>2</sup>  
José Márcio de MELLO<sup>2</sup>  
Henrique Ferrão SCOLFORO<sup>2</sup>  
Aliny Aparecida dos REIS<sup>2</sup>  
Lucas Rezende GOMIDE<sup>2</sup>

- RESUMO: O efeito decorrente da continuidade espacial entre parcelas vem sendo percebido desde, pelo menos, 1925, quando Fisher propôs os delineamentos experimentais, cujo intuito era neutralizar a influência desses efeitos. Porém, com avanço das técnicas espaciais, a presença desses efeitos deixou de ser vista como uma inconveniência a ser evitada, tornando-se uma fonte de informação valiosa. Diante disso, o presente estudo teve como objetivo analisar dados de um plantio experimental de candeia a partir de um modelo que considerasse o efeito da informação espacial, por meio de métodos geoestatísticos. O experimento foi instalado sob um delineamento em blocos casualizados composto por 4 blocos e 13 tipos de adubação, sendo avaliadas as variáveis diâmetro a 1,30 metros do solo (DAP) e a altura total (H) da árvore. Ao comparar os resultados alcançados a partir de uma análise clássica com os obtidos por meio do modelo espacial, observou-se que esse último foi mais eficaz. Assim, a agregação da componente espacial na análise de variância permitiu diferenciar os tratamentos de adubação para a variável DAP, na qual aplicação do adubo formulado NPK 8-28-16, cujo valor observado foi 6,0025 cm, propiciou uma maior condição de crescimento para a candeia, o que não foi percebido classicamente.
- PALAVRAS-CHAVE: Semivariograma; experimentação florestal; análise de variância.

<sup>1</sup>Universidade Federal de Lavras - UFLA, Departamento de Ciências Exatas, CEP: 37200-000, Lavras, MG, Brasil. E-mail: *cris.hnoqueira@yahoo.com.br*; *marcelo.oliveira@dex.ufla.br*

<sup>2</sup>Universidade Federal de Lavras - UFLA, Departamento de Ciências Florestais, CEP: 37200-000, Lavras, MG, Brasil. Email: *marcelufla@gmail.com*, *josemarcio@dcf.ufla.br*, *henriquescolforo@hotmail.com*, *alinyreis@hotmail.com*, *lucasgomide@dcf.ufla.br*

## 1 Introdução

A análise estatística de experimentos de campo é realizada, usualmente, por meio da análise de variância clássica, que, conforme destaca Banzatto e Kronka (2006), tem como uma das pressuposições básicas a independência dos erros, cujo alicerce encontra-se na aleatorização do experimento. Embora dados não pronunciados espacialmente tornem o processo de análise mais simples e conveniente por parte do pesquisador, Cressie (1993) argumenta que a desconsideração dos efeitos espaciais pode impedir que reais diferenças entre os tratamentos analisados sejam levantadas. Assim, Richter e Kroschewski (2012) afirmam que a probabilidade de detectar diferenças de tratamentos muitas vezes pode ser aumentada usando geoestatística em vez de modelos estatísticos clássicos.

Além disso, conforme afirmam Stroup, Baenziger e Muiltze (1994), as análises convencionais não neutralizam de maneira eficiente os efeitos da variabilidade espacial em muitos experimentos. Essa ineficiência, segundo Casanoves, Macchiavelli e Balzarini (2005), é percebida principalmente em experimentos com um número elevado de tratamentos ou quando existe uma elevada variabilidade no material experimental, a qual resulta em uma perda de precisão na análise.

Nesse sentido, “diversos estudos relacionados a fenômenos naturais que ocorrem em solos, atmosfera, corpos d’água, entre outras estruturas físicas espaciais, apontaram para a autocorrelação espacial dos erros”(MAIA et al., 2013). Mais especificamente, Silva, Guimarães e Pedrosa (2004) avaliaram o efeito da dependência espacial sobre a seleção de progênies de milho, considerando na análise dos dados o modelo com erros independentes e com erros espacialmente autocorrelacionados. Esses autores verificaram que o modelo com erros espacialmente correlacionados ajustou-se melhor aos dados, apresentando um maior poder de distinção e seleção das progênies avaliadas.

Nesse mesmo contexto, avaliando a existência de correlação espacial em um experimento instalado sob um delineamento em blocos casualizados (DBC), Reis e Miranda Filho (2003) chegaram a conclusão de que a modelagem espacial foi mais apropriada, resultando na redução das estimativas das variâncias residuais e, conseqüentemente, em melhoria nos ganhos esperados pela seleção dos tratamentos. Desta forma, entende-se que a incorporação da estrutura de dependência espacial na análise de variância tornou-se uma metodologia importante e necessária na avaliação de experimentos de campo.

Devido a isso, o presente estudo teve como objetivo principal a análise de dados de um experimento de candeia (*Eremanthus erythropappus* (DC.) McLeish), no qual foi avaliada a influência da adubação no diâmetro e na altura da planta, a partir de um modelo que considerasse o efeito de dependência espacial. Havendo diferenças significativas entre os tratamentos analisados, conduziu-se à aplicação de um teste de Tukey aproximado, como forma de discernir a influência média dos tipos de adubação. Além disso, no intuito de verificar os ganhos efetivos dessa abordagem, comparou-se os resultados obtidos por meio da análise de variância considerando o modelo com autocorrelação espacial dos erros com aqueles obtidos pela análise de variância clássica.

## 2 Materiais e métodos

### 2.1 Descrição da área de estudo

A área de estudo está localizada no município de Baependi, região sul do estado de Minas Gerais, cujas coordenadas geográficas são 21° 58' 33" de latitude Sul e 44° 44' 36" de longitude Oeste com 1500 m acima do nível do mar.

O clima da região, segundo a classificação Thorntwaite, é B2-úmido, mesotérmico úmido com verões quentes e chuvosos, e invernos frios e secos. A temperatura média anual varia entre 18 °C e 19 °C e a média anual de precipitação é de 1400 mm. Os meses mais chuvosos correspondem a dezembro, janeiro e fevereiro, e as menores precipitações ocorrem em junho, julho e agosto.

O experimento foi instalado em 2004, numa área de 1,5 ha, disposto em delineamento de blocos casualizados (DBC), composto por 4 blocos e 13 tratamentos. Cada unidade experimental continha 50 plantas úteis medidas e 4 plantas de bordadura, espaçadas em 2,0 x 2,5 m.

Os tratamentos desse experimento referem-se aos tipos de adubação aplicados em cada unidade experimental, identificados conforme apresentado na Tabela 1.

Tabela 1 - Tratamentos de fertilização avaliados no experimento

Tratamento	Adubação
1	testemunha absoluta
2	completo: Ca + NPK + micros
3	N: Ca + PK + micros
4	P <sub>2</sub> O <sub>5</sub> : Ca + NK + micros
5	K <sub>2</sub> O: Ca + NP + micros
6	micros: Ca + NPK
7	calcário: PK + micros
8	completo + adubação orgânica
9	testemunha absoluta + adubação orgânica
10	adubo formulado NPK (8-28-16)
11	completo 2: Ca + gesso + NPK + micros
12	completo 3: Ca + gesso + NPK* + micros
13	completo 4: Ca + adubo formulado NPK (6-30-6) + NK + micros

\* (100g de superfosfato simples + 200g de fosfato reativa.)

A aplicação dos elementos indexados nos devidos tratamentos referem-se às seguintes quantidades: N = 100 g/cova de Sulfato de Amônia, P = 200 g/cova de Superfosfato Simples, K = 50 g/cova de Cloreto de Potássio, Ca = 200 g/cova de Calcário, Gesso = 100 g/cova, micros = 0,5 g/cova de micronutrientes, adubo orgânico = 10 L/cova, adubo formulado NPK (8-28-16) = 150 g/cova, adubo formulado NPK (6-30-6) = 100 g/cova. Para o tratamento 12 em especial aplicou-se P = 100 g/cova Superfosfato Simples mais + 200g/cova de Fósforo Reativo. Os micronutrientes foram compostos pelos seguintes

elementos: Zn (9,0%) + B (1,8%) + Cu (0,8%) + Fe (3,0%) + Mn (2,0%) + Mo (0,1%).

A coleta dos dados foi realizada no mês junho do ano de 2012, onde foram coletados os diâmetros a 1,30 metros do solo (DAP) em centímetros (cm) e as alturas (H) em metros (m) de todos os indivíduos pertencentes ao experimento. Para análise espacial das variáveis DAP e H foram utilizados valores médios referentes a cada unidade experimental, cuja coordenada geográfica refere-se ao vértice inferior da parcela.

## 2.2 Semivariograma

As semivariâncias foram obtidas a partir do estimador de Matheron que, segundo Cressie (1993), pode ser definido por meio da seguinte expressão:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} [z(x_i) - z(x_i + \mathbf{h})]^2, \quad (1)$$

em que

$\hat{\gamma}(\mathbf{h})$  é o estimador de semivariância;

$n(\mathbf{h})$  representa o número de pares de valores medidos separados pelo vetor de distâncias euclidianas  $\mathbf{h}$ ;

$z(x_i)$  e  $z(x_i + \mathbf{h})$  são realizações da variável aleatória  $Z$ , nas coordenadas  $x_i$  e  $x_i + \mathbf{h}$ , de tal modo que esses pontos estão separados pela distância  $\mathbf{h}$ .

Na busca por uma melhor descrição do comportamento espacial, ajustou-se aos semivariogramas experimentais os modelos esférico, exponencial e gaussiano, cujas funções de semivariância são definidas como:

*Modelo esférico:*

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , \text{ se } \mathbf{h} = 0; \\ \tau^2 + \varphi^2 \left[ \frac{3}{2} \left( \frac{\mathbf{h}}{\phi} \right) - \frac{1}{2} \left( \frac{\mathbf{h}}{\phi} \right)^3 \right] & , \text{ se } 0 < \mathbf{h} < \phi; \\ \tau^2 + \varphi^2 & , \text{ se } \mathbf{h} \geq \phi; \end{cases} \quad (2)$$

*Modelo exponencial:*

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , \text{ se } \mathbf{h} = 0; \\ \tau^2 + \varphi^2 \left[ 1 - \exp \left( \frac{-3\mathbf{h}}{\phi} \right) \right] & , \text{ se } \mathbf{h} \neq 0; \end{cases} \quad (3)$$

*Modelo gaussiano:*

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , \text{ se } \mathbf{h} = 0; \\ \tau^2 + \varphi^2 \left[ 1 - \exp \left[ - \left( \frac{\sqrt{3}\mathbf{h}}{\phi} \right)^2 \right] \right] & , \text{ se } \mathbf{h} \neq 0. \end{cases} \quad (4)$$

em que o alcance ( $\phi$ ) é distância dentro da qual as amostras apresentam-se correlacionadas espacialmente, o efeito pepita ( $\tau^2$ ) representa a descontinuidade do semivariograma para distâncias menores do que a menor distância observada na amostra, sendo considerada uma variação totalmente aleatória dos dados e, por fim, a contribuição ( $\varphi^2$ ) é o valor associado à variância estrutural, ou seja, à variabilidade que pode ser explicada por um modelo com estrutura de dependência espacial. Esses modelos foram ajustados por meio dos métodos de quadrados mínimos ordinários e de máxima verossimilhança.

Na abordagem geoestatística, o ajuste do semivariograma exerce um papel fundamental e determinante para a eficiência dessa análise. Diante disso, a modelagem do semivariograma constituiu uma etapa mais cautelosa desse trabalho, na qual a escolha pelo melhor modelo ajustado deu-se pelo método de validação cruzada. Basicamente, conforme salienta Isaaks e Srivastava (1989), essa técnica consiste em excluir uma observação e pelas observações remanescentes estimar a observação excluída.

Com base na validação cruzada, a comparação entre os métodos de ajuste e os modelos avaliados foi realizada por meio do Erro Médio Reduzido ( $\overline{ER}$ ) e do Desvio Padrão do Erro Reduzido ( $S_{ER}$ ) que, segundo McBratney e Webster (1986), são dados, respectivamente por:

$$\overline{ER} = \frac{1}{n} \sum_{i=1}^N \left( \frac{Z(x_i) - \hat{Z}(x_i)}{\sigma_{x_i}} \right) \quad (5)$$

e

$$S_{ER} = \sqrt{\frac{1}{n} \sum_{i=1}^N \left( \frac{Z(x_i) - \hat{Z}(x_i)}{\sigma_{x_i}} \right)^2} \quad (6)$$

em que

$Z(x_i)$  é o valor observado no ponto  $x_i$ ;

$\hat{Z}(x_i)$  é o valor estimado para o ponto  $x_i$ ;

$\sigma_{x_i}$  o desvio padrão da krigagem no ponto  $x_i$  e

$n$  é o número de observações.

Segundo Cressie (1993), o Erro Médio Reduzido mais próximo de zero e o Desvio Padrão do Erro Reduzido mais próximo de um são critérios para escolher o melhor modelo de semivariância ajustado. Vale ressaltar que todas as análises estatísticas foram executadas com o auxílio do *software* R (R DEVELOPMENT CORE TEAM, 2014).

### 2.3 Modelagem linear dos dados

Os dados foram analisados segundo um delineamento em blocos ao acaso, cujo modelo estatístico pode ser representado por:

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad (7)$$

sendo que

$y_{ij}$  é a observação do  $i$ -ésimo tratamento do  $j$ -ésimo bloco, com  $i = 1, 2, \dots, t$  e  $j = 1, 2, \dots, r$ ;

$\mu$  é uma constante inerente a cada observação;

$\alpha_i$  é o efeito do  $i$ -ésimo tratamento;

$\beta_j$  é o efeito relacionado ao bloco  $j$ ;

$e_{ij}$  é o erro aleatório associado à  $ij$ -ésima observação.

O modelo linear geral dos dados, em forma matricial, pode ser descrito como:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \quad (8)$$

em que

$\mathbf{Y}$  é o vetor de observações das variáveis de ordem  $n \times 1$ , sendo  $n = tr$ ;

$\mathbf{X}$  é a matriz de incidência dos parâmetros do modelo de ordem  $n \times p$ , em que  $p = t + r + 1$ ;

$\boldsymbol{\theta}$  é o vetor de parâmetros do modelo de ordem  $p \times 1$ ;

$\boldsymbol{\varepsilon}$  é o vetor de erros aleatórios de ordem  $n \times 1$ , sendo usual assumir que  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ .

Os dados foram analisados por meio da aplicação de dois modelos, o primeiro considerando a estrutura de dependência espacial entre os mesmos e o segundo assumindo a independência dos erros. Assim, foram consideradas as seguintes estruturas para a matriz de covariância:

i.  $\boldsymbol{\Sigma} = \mathbf{V}\sigma^2$  ( $\mathbf{V} \neq \mathbf{I}$  é uma matriz de correlação),

ii.  $\boldsymbol{\Sigma} = \mathbf{I}\sigma^2$  ( $\mathbf{I}$  é uma matriz identidade).

Ao considerar os erros espacialmente dependentes, a matriz  $\Sigma$  foi obtida a partir do ajuste do semivariograma. Por meio das relações  $C(\mathbf{h}) = C(\mathbf{0}) - \gamma(\mathbf{h})$  e  $\rho(\mathbf{h}) = C(\mathbf{h})/C(\mathbf{0})$ , em que  $C(\mathbf{0}) = \sigma^2$  e  $C(\mathbf{h})$  e  $\rho(\mathbf{h})$  denotam, respectivamente, a covariância e a correlação na distância  $\mathbf{h}$ , é possível reescrever  $\Sigma$  como

$$\Sigma = \mathbf{V}\sigma^2 = \begin{bmatrix} 1 & \rho(\mathbf{h}_{12}) & \cdots & \rho(\mathbf{h}_{1n}) \\ \rho(\mathbf{h}_{12}) & 1 & \cdots & \rho(\mathbf{h}_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(\mathbf{h}_{1n}) & \rho(\mathbf{h}_{2n}) & \cdots & 1 \end{bmatrix} \sigma^2;$$

de modo que  $\sigma^2 = \tau^2 + \varphi^2$  corresponde ao patamar do semivariograma e  $\mathbf{h}_{ij}$  é o vetor de distância entre os erros  $e_i$  e  $e_j$ , respectivamente alocados nas posições  $x_i$  e  $x_j$ , ou seja,  $\mathbf{h}_{ij} = \|x_i - x_j\|$ .

Os parâmetros do modelo considerando os erros espacialmente dependentes foram obtidos pelo método de quadrados mínimos generalizados é dado por

$$\boldsymbol{\theta}^0 = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} \quad (9)$$

sendo que a notação  $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$  indica uma inversa generalizada da matriz  $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})$  e  $\boldsymbol{\theta}^0$  uma das infinitas soluções existentes.

No intuito de melhorar a eficiência das estimativas dos parâmetros adotou-se um processo iterativo para a obtenção de  $\Sigma$ , conforme proposto por Pontes e Oliveira (2004) e utilizado por Nogueira (2013).

Para obtenção das somas de quadrados da análise de variância considerou-se uma partição, mencionada por Searle (1987), do modelo apresentado em (8) da seguinte maneira:

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\mu} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (10)$$

sendo  $\mathbf{Y}$  o vetor de observações,  $\boldsymbol{\mu}$  uma constante inerente a cada observação,  $\boldsymbol{\alpha}$  um vetor  $t \times 1$  que contém os efeitos de tratamento,  $\boldsymbol{\beta}$  um vetor  $r \times 1$  contendo os efeitos de bloco,  $\mathbf{X}_1$  um vetor de uns,  $\mathbf{X}_2$  a matriz de incidência dos efeitos de tratamento de ordem  $n \times t$ ,  $\mathbf{X}_3$  a matriz de incidência dos efeitos de bloco de ordem  $n \times r$ , de tal modo que  $\mathbf{X} = [\mathbf{X}_1 \dot{\vdots} \mathbf{X}_2 \dot{\vdots} \mathbf{X}_3]$  e  $\boldsymbol{\theta}' = [\boldsymbol{\mu}' \dot{\vdots} \boldsymbol{\alpha}' \dot{\vdots} \boldsymbol{\beta}']$  e, por fim,  $\boldsymbol{\varepsilon}$  representa um vetor de erros.

Definindo  $r(\mathbf{A})$  como o *rank* da matriz  $\mathbf{A}$ , Nogueira (2013) utilizou essa partição para mostrar que a análise de variância considerando a dependência espacial dos erros de um delineamento em blocos casualizados, mostrada na Tabela 2.

Tabela 2 - Análise de variância de um DBC, considerando erros dependentes

FV	GL	SQ
Tratamento	$r(\mathbf{X}_2) - r(\mathbf{X}_1)$	$\mathbf{Y}'\mathbf{P}_2\mathbf{Y}$
Bloco	$r(\mathbf{X}) - r(\mathbf{X}_2)$	$\mathbf{Y}'(\mathbf{P} - \mathbf{P}_1 - \mathbf{P}_2)\mathbf{Y}$
Resíduo	$n - r(\mathbf{X})$	$\mathbf{Y}'(\mathbf{V}^{-1} - \mathbf{P})\mathbf{Y}$
Total	$n - r(\mathbf{X}_1)$	$\mathbf{Y}'(\mathbf{V}^{-1} - \mathbf{P}_1)\mathbf{Y}$

em que:

$$\mathbf{P} = \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}; \quad \mathbf{P}_1 = \mathbf{V}^{-1}\mathbf{X}_1(\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{V}^{-1};$$

$$\mathbf{P}_2 = \mathbf{R}\mathbf{X}_2(\mathbf{X}_2'\mathbf{R}\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{R}_v \quad \text{e} \quad \mathbf{R} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}_1(\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{V}^{-1}$$

#### 2.4 Teste de comparação de médias

Para as análises que apontaram diferença significativa entre as médias dos tratamentos, aplicou-se o teste de Tukey com um nível de 5% de probabilidade, cujo objetivo foi discriminar para quais tratamentos de adubação o plantio de candeia apresentou maior desenvolvimento.

Esse teste tem o intuito de comparar qualquer contraste entre duas médias de tratamentos. De acordo com Pimentel-Gomes (2009), o teste de Tukey declara duas médias,  $\bar{y}_u$  e  $\bar{y}_v$ , significativamente diferentes se o valor absoluto de suas diferenças amostrais ultrapassar

$$\Delta = q_{(\alpha,t,\psi)} \sqrt{\frac{1}{2} \hat{V}ar(\bar{y}_u - \bar{y}_v)},$$

em que  $\hat{V}ar(\cdot)$  é uma função que representa a variância estimada e  $q_{(\alpha,t,\psi)}$  é o valor da amplitude total estudentizada a um nível  $\alpha\%$  de probabilidade, em um experimento com  $t$  tratamentos e grau de liberdade do resíduo igual a  $\psi$ .

Assim, rejeita-se a igualdade da média de dois tratamentos se  $|\bar{y}_u - \bar{y}_v| > \Delta$ , em que  $\bar{y}_u$  e  $\bar{y}_v$  representam as médias das observações que receberam, respectivamente, o  $u$ -ésimo e  $v$ -ésimo tratamento.

Segundo Richter et al. (2014), a consideração de qualquer correlação espacial existente como um componente do modelo pode ser benéfico na análise dos dados, porém, requer a seleção de um modelo espacial adequado e modificação de testes clássicos de contrastes de tratamento

Como a análise espacial considera a covariância não nula, a variância da diferença entre as médias de dois tratamentos  $u$  e  $v$ , nesse caso, é desdobrada em:



$$\begin{aligned}
Var(\bar{y}_u. - \bar{y}_v.) &= Var(\bar{y}_u.) + Var(\bar{y}_v.) - 2Cov(\bar{y}_u., \bar{y}_v.) \\
&= \frac{1}{r^2} Var\left(\sum_{j=1}^r y_{uj}\right) + \frac{1}{r^2} Var\left(\sum_{j=1}^r y_{vj}\right) - \frac{2}{r^2} Cov\left(\sum_{j=1}^r y_{uj}, \sum_{k=1}^r y_{vk}\right) \\
&= \frac{1}{r^2} \sum_{j=1}^r \sum_{k=1}^r Cov(e_{uj}, e_{uk}) + \frac{1}{r^2} \sum_{j=1}^r \sum_{k=1}^r Cov(e_{vj}, e_{vk}) - \frac{2}{r^2} \sum_{j=1}^r \sum_{k=1}^r Cov(e_{uj}, e_{vk})
\end{aligned}$$

A utilização de  $q_{(\alpha,t,\psi)}$  no cálculo de  $\Delta$ , quando a correlação espacial está presente, resulta em um teste de comparação entre médias aproximado.

### 3 Resultados e discussões

Primeiramente, foram calculados os vetores de erros considerando as variáveis DAP e H, cujos parâmetros foram estimados pela maneira clássica, considerando  $\mathbf{V} = \mathbf{I}$  uma matriz identidade. A partir daí, no intuito de verificar a semelhança desses erros em função da distância que os separam, construiu-se os gráficos, apresentado na Figura 1, no qual é possível observar a distribuição espacial das observações de acordo com o quartil ao qual pertence, de modo que os maiores valores do erro analisado encontram-se nos maiores quartis.

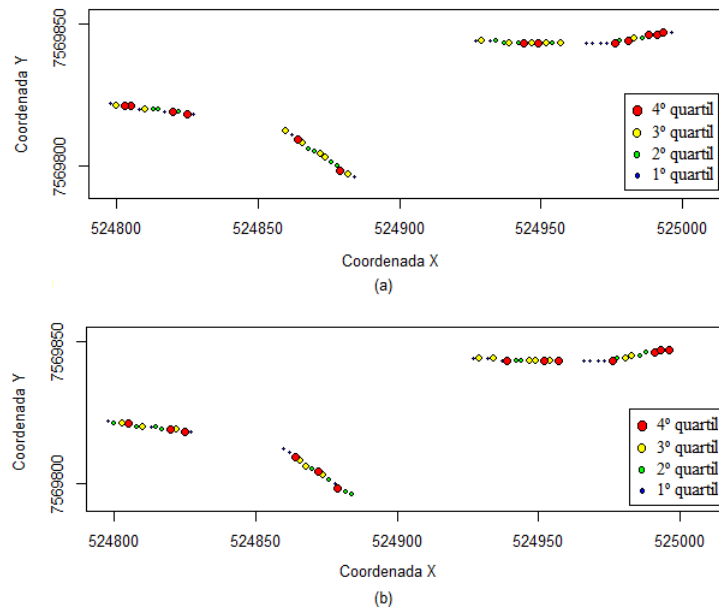


Figura 1 - Grid com os quartis dos erros referentes à variável DAP (a) e H (b).

A partir desses erros, prosseguiu-se na construção do semivariograma experimental, sendo que, considerando o erro como uma variável isotrópica, construiu-se o semivariograma *omnidirecional*, de forma que, para o cálculo de suas semivariâncias, foram consideradas apenas as distâncias, em qualquer direção, entre dois pontos amostrados.

Com base nesse semivariograma ajustou-se os modelos esférico, exponencial e gaussiano pelos métodos de quadrados mínimos ordinários (OLS) e de máxima verossimilhança (ML). A seleção do modelo e método de ajuste consistiu em empregar a técnica de validação cruzada para cada combinação dos modelos e métodos analisados. A seguir, estão apresentados na Tabela 3, as estatísticas referentes ao erro médio reduzido ( $\overline{ER}$ ) e o desvio padrão do erro reduzido ( $S_{ER}$ ) obtidos mediante a técnica de validações cruzada, para cada configuração considerada.

Tabela 3 - Estatísticas de validação cruzada para as variáveis DAP e H

Variável	Método	Modelo	$\overline{ER}$	$S_{ER}$
DAP	OLS	Esférico	$9,3675 * 10^{-17}$	2,3149
		Exponencial	$1,9402 * 10^{-16}$	2,2419
		Gaussiano	$3,1092 * 10^{-17}$	2,3937
	ML	Esférico	$-8,2908 * 10^{-17}$	2,4447
		Exponencial	$-8,6603 * 10^{-17}$	2,3675
		Gaussiano	$-2,5460 * 10^{-16}$	3,9539
H	OLS	Esférico	$-2,4980 * 10^{-16}$	8,4888
		Exponencial	$5,0174 * 10^{-17}$	6,0980
		Gaussiano	$-1,3883 * 10^{-15}$	8,1268
	ML	Esférico	$1,5266 * 10^{-16}$	9,3011
		Exponencial	$2,1372 * 10^{-15}$	8,9213
		Gaussiano	$-9,0479 * 10^{-16}$	9,3475

Analisando os resultados apresentados na Tabela 3, verificou-se que em todas as combinações o erro médio reduzido foi muito próximo de zero, condição desejada na escolha pelo melhor ajuste. Assim, analisando o desvio padrão do erro reduzido observa-se que os modelos ajustados pelo método de quadrados mínimos ordinários apresentaram melhor desempenho para as duas variáveis em estudo.

Além disso, o desvio padrão do erro reduzido foi mais próximo de um quando considerado o modelo exponencial, tanto para a variável DAP quanto para a H, optando-se pela escolha desse modelo. Na Tabela 4 são apresentados os valores dos parâmetros estimados para o modelo exponencial, a partir do método de quadrados mínimos, e estabilizados pelo processo iterativo. Na Figura 2 é possível visualizar a curva ajustada desse modelo para as duas variáveis analisadas. A partir do semivariograma experimental, é possível observar alguns pontos de semivariância discrepantes ocasionado, possivelmente, pela presença de *outliers* no conjunto de dados.

Tabela 4 - Parâmetros estimados do semivariograma para as variáveis DAP e H

Variável	Parâmetros	Valores
DAP	$\sigma^2$	0,3118
	$\tau^2$	0,0000
	$\phi$	3,8416
H	$\sigma^2$	0,0402
	$\tau^2$	0,0000
	$\phi$	5,1791

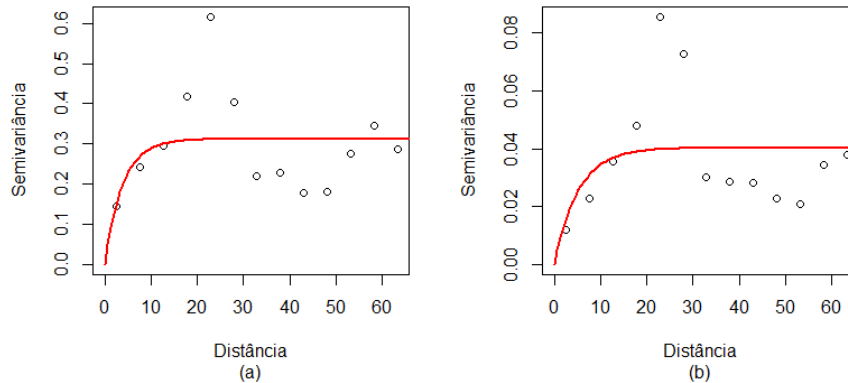


Figura 2 - Modelo de semivariograma ajustado aos erros referentes à variável DAP (a) e H (b).

Com base nos modelos ajustados aos dados, foi possível verificar que, para a variável DAP, as parcelas separadas por uma distância menor que 3,8416 m são espacialmente dependentes. Já para a variável H, observa-se que parcelas separadas por uma distância maior que 5,1791 m não apresentam dependência espacial.

A partir do ajuste desses modelos obteve-se a matriz de covariância dos erros ( $\Sigma$ ) para ambas variáveis e, a partir desta, foi obtida a matriz de correlação ( $V$ ). De posse da matriz de correlação, procedeu-se o cálculo da análise de variância com erros espacialmente correlacionados conforme descrito na Tabela 2.

A Tabela 5 apresenta a análise de variância para a variável DAP referente ao modelo com erros correlacionados espacialmente. De acordo com esta análise, pode-se concluir que os efeitos de tratamento, referentes à adubação, foram significativos ao nível de 5% de probabilidade e, com isso, que as médias dos tratamentos diferem estatisticamente entre si.

Tabela 5 - Análise de variância, considerando a análise espacial, para a variável DAP

FV	GL	SQ	QM	$F_0$	valor p
Tratamento	12	12,7743	1,0645	2,6286	0,0124
Blocos	3	14,7038	4,9013	12,1027	0,0000
Resíduo	36	14,5792	0,4050		
Total	51	42,0573			

No intuito de discriminar os diferentes tratamentos, aplicou-se o teste aproximado de comparação entre médias de Tukey. Nesse teste, as covariâncias estimadas foram obtidas com base na estimativas de semivariância. A Tabela 6 apresenta os resultados desse teste com os respectivos ordenamentos das médias e tratamentos, bem como a discriminação dos tratamentos pelo teste sendo que, segundo Banzatto e Kronka (2006), médias seguidas de mesma letra não diferem pelo teste de Tukey.

Tabela 6 - Comparação entre médias dos diferentes tratamentos de fertilização para a variável DAP

Tratamento	$DAP_{medio}$	Grupo
10	6,0025	a
2	5,5575	ab
9	5,4550	ab
6	5,4325	ab
7	5,4075	ab
13	5,3750	ab
8	5,3525	ab
11	5,2600	ab
4	5,0575	ab
3	5,0350	ab
12	4,9450	ab
5	4,8500	ab
1	4,4325	b

A partir desse teste foi possível diagnosticar que a média proveniente do tratamento 10 é significamente superior à média oriunda do tratamento 1 (testemunha absoluta). Assim, pode-se concluir que o maior ganho em crescimento do diâmetro a 1,30 metros do solo (DAP) ocorreu quando foi aplicado o adubo formulado NPK 8-28-16.

Já na Tabela 7 encontra-se a análise de variância considerando os erros correlacionados para a variável H. A partir desta, pode-se observar que mesmo

com a inclusão da correlação espacial na análise de variância, os tratamentos não apresentaram diferença significativa, ao nível de significância de 5% de probabilidade.

Tabela 7 - Análise de variância, considerando a análise espacial, para a variável H

FV	GL	SQ	QM	$F_0$	valor p
Tratamento	12	0,8762	0,0730	1,6718	0,1152
Blocos	3	4,3441	1,4480	33,1350	0,0000
Resíduo	36	1,5724	0,0437		
Total	51	6,7927			

No intuito de verificar a maior eficiência da modelagem com o acréscimo da informação espacial, apresenta-se nas Tabelas 8 e 9 os resultados obtidos por meio da análise de variância clássica cujos erros são considerados independentes, respectivamente, para as variáveis DAP e H.

Tabela 8 - Análise de variância, considerando a análise clássica, para a variável DAP

FV	GL	SQ	QM	$F_0$	valor p
Tratamento	12	7,1648	0,5971	1,8757	0,0721
Blocos	3	43,4491	14,4830	45,4996	0,0000
Resíduo	36	11,4592	0,3183		
Total	51	62,0731			

Tabela 9 - Análise de variância, considerando a análise clássica, para a variável H

FV	GL	SQ	QM	$F_0$	valor p
Tratamento	12	0,7604	0,0634	1,6704	0,1156
Blocos	3	14,4186	4,8062	126,6912	0,0000
Resíduo	36	1,3657	0,0379		
Total	51	16,5447			

Ao observar essas análises, como o *valor p* foi menor que 0,05, pode-se concluir que não existem evidências suficientes, ao nível de 5% de probabilidade, para considerar que os efeitos de adubação interferiram de forma diferente nas variáveis resposta altura e diâmetro da árvore.

Analisando o fator bloco em todas as análises realizadas neste trabalho, verifica-se que as diferenças entre as médias dos blocos são altamente significativas,

condizente com o que se esperava no momento de instalação do experimento, ou seja, que a blocagem realmente colaborasse para o controle do efeito do ambiente no crescimento das plantas de candeia.

Comparando-se os resultados entre a análise espacial e a clássica para a variável DAP, Tabelas 5 e 8, foi notável o ganho a partir da abordagem geoestatística durante a análise estatística no experimento, reduzindo o *valor p* de 0,0721 para 0,0124, passando, assim, a detectar diferenças significativas entre as médias dos tratamentos, considerando um nível de 5% de probabilidade.

Além disso, avaliando os resultados alcançados para a variável DAP, pode-se inferir que a não consideração da dependência espacial dos erros ocasionou uma maior probabilidade de ocorrência do erro tipo II, no qual tratamentos diferentes na prática apontam-se estatisticamente semelhantes.

Segundo Legendre et al., (2002), este tipo de erro ocorre pela alta proximidade entre as parcelas e, conseqüentemente, a existência da correlação espacial, que por sua vez não foi contornada de forma eficaz pela aleatorização das unidades experimentais.

Esses resultados condizem com Scolforo et al. (2012) que, para a mesma área de estudo considerando a presença da correlação espacial a partir da modelagem autorregressiva, encontraram melhoria nos resultados ao observar a variável DAP. Para a variável H, os autores também não encontraram efeito significativo do espaço ao se utilizar esse tipo de abordagem. Nesse mesmo âmbito, Grondona e Cressie (1991) e Nogueira (2013), nos quais, ao trabalharem com dados simulados instalados sob um delineamento em blocos casualizados, concluíram que a abordagem espacial se mostrou mais eficiente em relação a abordagem clássica.

Na comparação entre as Tabelas 7 e 9 esta eficiência não foi, de forma tão intensa, verificada para a variável altura. Isso pode ser justificado devido ao fato que, na idade considerada, já ocorre uma diminuição considerável da competição entre as plantas. Assim, o efeito da adubação é mais pronunciado no diâmetro. Isto faz com que haja diferença de ganhos em crescimento desta característica em função dos benefícios da adubação. Porém, acredita-se que a análise espacial continue sendo a mais adequada, uma vez que não viola nenhum tipo de suposição pré-estabelecida no modelo, sendo, do ponto de vista da modelagem científica, superior, uma vez que considera a correlação espacial atuante no fenômeno.

## Conclusões

Diante do que foi exposto, entende-se que este procedimento de análise espacial tem uma maior capacidade discriminatória dos tratamentos, uma vez que possibilitou identificar diferenças para o DAP quanto aos tratamentos quando a componente espacial foi agregada.

Com isso, a identificação da presença de correlação espacial no ambiente faz-se necessária para que não haja o enviesamento da informação durante as análises estatísticas e, com isso, alterando as conclusões obtidas acerca do experimento.

Para a variável DAP, a aplicação de NPK 8-26-16, cujo valor médio observado foi de 6,0025 cm, foi considerada superior à testemunha geral, cujo valor médio

encontrado foi 4,4325 cm. Sendo assim, recomenda-se a utilização do adubo formulado NPK 8-28-16 na implantação de candeais na região sul do estado de Minas Gerais para a obtenção de um maior crescimento diamétrico da planta.

## Agradecimentos

Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo apoio financeiro.

NOGUEIRA, C. H.; OLIVEIRA, M. S; MELLO, J. M; RAIMUNDO, M. R; SCOLFORO, H. F; REIS, A. A; GOMIDE, L. R. Spatial modeling in the analysis of a experimental planting candeia. *Rev. Bras. Biom.*, São Paulo, v.33, n.1, p.14-29, 2015.

■ **ABSTRACT:** *With the advancement of space technology, the presence of spatial effects becomes a source of information that can result in more accurate analysis. Thus, the present study aimed to analyze data from an experimental planting of candeia from a model that considered the effect of spatial information by geostatistical methods. The experiment was conducted under a randomized block design consisting of four blocks and thirteen types of fertilization. The variables diameter at breast height (DBH) and total height (H) of the trees were evaluated. By comparing the results obtained from a standard analysis, which is based on the assumption of independent errors, with those obtained by the spatial model, we observed that the second approach was more effective in detecting difference between treatment means. Thus, the aggregation of spatial component in the analysis of variance allowed the differentiation of the fertilization treatments for variable diameter. Therefore the application of formulated fertilizer NPK 8-28-16 provided the largest growth.*

■ **KEYWORDS:** *Semivariogram; forest experiment; analysis of variance.*

## Referências

BANZATTO, D. A.; KRONKA, S. N.; *Experimentação agrícola*. 4.ed. Jaboticabal: Funep, 2006. 237p.

CASANOVES, F.; MACCHIAVELLI, R.; BALZARINI, M. Error Variation in multienvironment peanut trials: Within-trial spatial correlation and between-trial heterogeneity. *Crop Science*, Madison, v.45, p.1927-1933, 2005.

CRESSIE, N. *Statistics for spatial data*. 2<sup>nd</sup> ed. New York: J. Wiley, 1993. 900 p.

GRONDONA, M. O.; CRESSIE, N. *Using spatial considerations in the analysis of experiments*. *Technometrics*, Alexandria, v.33, n.4, p.381-392, 1991.

ISAASKS, E. H; SRIVASTAVA, R. M. *Applied geostatistics*. Oxford: Oxfors University, 1989. 560p.

LEGENDRE, P.; DALE, M. R. T.; FORTIN, M. J.; GUREVITCH, J.; HOHN, M.; MYERS, D. The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, n.25, p.601-615, 2002.

- MAIA, E.; SIQUEIRA, D. L.; CARVALHO, S. A.; PETERNELLI, L. A.; LATADO, R. R. Aplicação da análise espacial na avaliação de experimentos de seleção de clones de laranja Pêra. *Ciência Rural*, Santa Maria, v.43, n.1, p.8-14, 2013.
- MCBRATNEY, A. B.; WEBSTER, R. Choosing functions for semivariograms of soil properties and fitting them to sampling estimates. *Journal of Soil Science*, New Brunswick, v.37, n.6, p.617-639, 1986.
- NOGUEIRA, C. H. *Análise de variância com dependência espacial sob uma abordagem geoestatística*. 2013. 123 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, 2013.
- PIMENTEL-GOMES, F. *Curso de estatística experimental*. 15.ed. Piracicaba: FEALQ, 2009, 451p.
- PONTES, J. M.; OLIVEIRA, M. S. Uma proposta alternativa para a análise de experimentos de campo utilizando a geoestatística. *Ciênc. agrotec.* v.28, n.1, p.135-141, 2004.
- R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Version 3.1.0. Vienna: R Foundation for Statistical Computing, 2014. Disponível em: <<http://www.r-project.org>>. Acesso em: 10 mai. 2014.
- REIS, A. J. dos S.; MIRANDA FILHO, J. B. de. Autocorrelação espacial na avaliação de compostos de milho para resistência à lagarta do cartucho (*Spodoptera frugiperda*). *Pesquisa Agropecuária Tropical*, Goiânia, v.33 n.2, p.65-72, 2003.
- RICHTER, C.; KROSCHEWSKI, B. Geostatistical Models in Agricultural Field Experiments: Investigations Based on Uniformity Trials. *Agronomy Journal*, Madison, v.104, n.1, p.91-105, 2012.
- RICHTER, C.H.; KROSCHEWSKI, B.; PIEPHO, H.-P.; SPIELKE, J. Treatment comparisons in agricultural field trials accounting for spatial correlation. *Journal of Agricultural Science*, Cambridge University Press, v.1, p.1-21, 2014.
- SCOLFORO, J. R. S.; OLIVEIRA, A. D. de ; DAVIDE, A. C. *Manejo Sustentável da Candeia: o caminhar de uma nova experiência florestal em Minas Gerais*. 1.ed. Lavras: Editora UFLA, 2012. v.1. 329p.
- SEARLE, S. R. *Linear models for unbalanced data*. New York: J. Wiley, 1987. 536p.
- SILVA, H. D.; GUIMARÃES, E. C.; PEDROSA, M. G. Incorporação da dependência espacial na análise de um experimento de avaliação de progênies de milho quanto à resistência à ferrugem comum. *Ciência e Agrotecnologia*, Lavras, v.28, n.5, p.1144-1150, 2004.
- STROUP, W. W.; BAENZIGER, P. S.; MULITZE, D. K. Removing spatial variation from wheat yield trials: a comparison of methods. *Crop Science*, Madison, v.34, p.62-66, 1994.

Recebido em 06.10.2014.

Approved after revised in 25.02.2015.