

SEÇÃO I - FÍSICA DO SOLO

SELEÇÃO DE MODELOS DE VARIABILIDADE ESPACIAL PARA ELABORAÇÃO DE MAPAS TEMÁTICOS DE ATRIBUTOS FÍSICOS DO SOLO E PRODUTIVIDADE DA SOJA⁽¹⁾

Mário Antonio Faraco⁽²⁾, Miguel Angel Uribe-Opazo⁽³⁾, Edson
Antonio Alves da Silva⁽⁴⁾, Jerry Adriani Johann⁽⁴⁾ & Joelmir
André Borssoi⁽⁴⁾

RESUMO

Pesquisas sobre a variabilidade espacial dos atributos do solo que influenciam a produtividade são de uma grande importância para o desenvolvimento de novas técnicas que beneficiam a agricultura. A variabilidade desses atributos pode ser avaliada por técnicas de geoestatística e auxiliar no mapeamento e manejo do solo. Este trabalho teve por objetivo avaliar a qualidade do ajuste dos modelos teóricos espaciais segundo o Critério de Informação de Akaike, de Filliben, de Validação Cruzada e o valor máximo do logaritmo da função verossimilhança, de dados da umidade do solo, da densidade do solo e da resistência do solo à penetração, nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m, e de produtividade da soja do ano agrícola 2004–2005. Os parâmetros dos modelos de variabilidade espacial foram estimados por meio dos métodos de mínimos quadrados ordinários, mínimos quadrados ponderados e máxima verossimilhança. A pesquisa foi desenvolvida em uma área de 57 ha de um Latossolo Vermelho distroférico, utilizando-se uma malha de 75 x 75 m georreferenciada. Concluiu-se que, dos métodos de avaliação de ajustes estudados, o da Validação Cruzada foi o mais adequado para escolha do melhor ajuste do modelo de variabilidade espacial; conseqüentemente têm-se mapas temáticos mais acurados.

Termos de indexação: geoestatística, métodos de estimação, validação do ajuste.

⁽¹⁾ Parte da Tese de Mestrado em Engenharia Agrícola com Área de Concentração em Sistemas Agroindustriais da Universidade Estadual do Oeste do Paraná – UNIOESTE. Recebido para publicação em fevereiro de 2007 e aprovado em outubro de 2007.

⁽²⁾ Mestre em Engenharia Agrícola pela Universidade Estadual do Oeste do Paraná – UNIOESTE. Rua Universitária 2069, CEP 85819-110 Cascavel (PR). E-mail: mafaraco@gmail.com

⁽³⁾ Professor Associado do Centro de Ciências Exatas e Tecnológicas – CCET, UNIOESTE. E-mail: mopazo@unioeste.br

⁽⁴⁾ Professor Assistente do Centro de Ciências Exatas e Tecnológicas – CCET, UNIOESTE. E-mails: edsonaasilva@gmail.com; jerryaj@pop.com.br; jborssoi@yahoo.com.br

SUMMARY: SELECTION CRITERIA OF SPATIAL VARIABILITY MODELS USED IN THEMATIC MAPS OF SOIL PHYSICAL ATTRIBUTES AND SOYBEAN YIELD

Studies on the spatial variability of soil attributes influencing crop productivity are important for the development of new technologies beneficial to agriculture. Geostatistical techniques can be used to evaluate the variability of soil attributes and contribute to soil mapping and management. The purpose of this paper was to evaluate the quality of the theoretical spatial model adjustments according to the Akaike Information and Filiben Criteria, Cross Validation and the maximum value of the log-likelihood function, of the soil humidity, of the soil density data and soil resistance to penetration, in the layers 0–0.1; 0.1–0.2; and 0.2–0.3 m and the soybean yield in the 2004–2005 growing season. The parameters of the spatial variability models were estimated by the methods of least ordinary squares, least weighted squares and maximum likelihood. The experiment was developed in an area of 57 ha with a regionally typical dystrophic Red Latosol (Oxisol). A spatially georeferenced 75 x 75 m regular mesh was used. Based on the results of the evaluation of adjustments it was concluded that the Cross Validation criterion was the most adequate to choose the best adjustment of the spatial variability model, resulting in more precise thematic maps.

Index terms: geostatistics, estimation methods, adjustment validation.

INTRODUÇÃO

A geoestatística surgiu para o estudo de variáveis regionalizadas, ou seja, o estudo de uma função espacial numérica, que varia de um local para outro, com continuidade aparente e cujos valores são relacionados com a posição espacial que ocupam. Oferece técnicas para elaboração de mapas do comportamento de variáveis georreferenciadas, utilizando o método de interpolação de informações a partir de dados obtidos em locais convenientemente amostrados e modelados em um semivariograma experimental.

Para modelar um conjunto de dados com uma estrutura de correlação, considera-se um processo estocástico gaussiano $\{Z(s), s \in S\}$, em que S é um subconjunto de R^d , sendo d um espaço euclidiano d -dimensional. Supõe-se que os dados $Z(s_1), \dots, Z(s_n)$ do processo são realizações conhecidas nos locais s_i , $i = 1, \dots, n$ em que s_i é um vetor d -dimensional ($d \geq 1$). Supõe-se que os dados, de modo geral, podem ser escritos como $Z(s_i) = \mu(s_i) + \epsilon(s_i)$, sendo $\mu(s_i)$ uma função determinística que descreve a componente estrutural de Z em s_i , que pode ser expressa como

$\mu(s_i) = \sum_{k=1}^p f_k(s) \beta_k$, sendo f_k uma função conhecida e β_k uma constante desconhecida a ser estimada, para $k = 1, \dots, p$ (caso particular $p = 1$, $\mu(s_i) = \beta_1$) e $\epsilon(s_i)$ é um termo estocástico, que varia localmente e depende espacialmente de $Z(s_i)$. Assume-se que o termo estocástico $\epsilon(s_i)$ tem média zero e a variação entre pontos no espaço é determinada pela função covariância $C(s_i, s_j) = \text{Cov}\{\epsilon(s_i), \epsilon(s_j)\}$ (Mardia & Marshall, 1984).

A semivariância é uma função da distância h , que é estimada em um conjunto discreto de distâncias (*lags*).

Em termos da covariância, a função semivariância define-se como $\gamma(h) = C(0) - C(h)$, sendo $h = ||s_i - s_j||$, $C(h) = C(s_i, s_j)$ e $C(0) = \sigma^2$ a variância das observações.

A partir da estimativa de $\gamma(h)$, ajusta-se um modelo que irá depender de parâmetros desconhecidos a serem determinados e com características espaciais. Segundo Isaaks & Srivastava (1989), a função semivariância $\gamma(h)$ é definida na equação 1 como:

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

em que $Z(s)$ é o valor da variável medida; $s \in S \subset R^d$, $d = 1, 2$ ou 3 , a localização; e h , a distância que separa duas amostras. Assim, o semivariograma experimental é um gráfico de dispersão de $\gamma(h)$ em função de h que permite uma análise variográfica do comportamento de $Z(s)$ (Cressie, 1993).

O principal estimador utilizado na construção do semivariograma experimental é o dos momentos, conhecido como semivariograma de Matheron, para processos estocásticos gaussianos. Para o caso de haver pontos discrepantes, que não podem ser eliminados ou substituídos, a literatura recomenda o estimador de Cressie & Hawkins (1980).

A partir da estimativa da semivariância empírica (ou experimental), ajusta-se um modelo teórico aos pontos obtidos. Escolher um modelo adequado é obter estimadores dos parâmetros efeito pepita C_0 , patamar $C_0 + C_1$ e alcance a , com métodos estatísticos de otimização como: mínimos quadrados ordinários, mínimos quadrados ponderados (Cressie, 1985) e máxima verossimilhança (Mardia & Marshall, 1984). Uma vez escolhido o modelo teórico de correlação espacial, resta saber se ele é eficiente para interpolar valores, permitindo estimativas confiáveis para construção de mapas temáticos (Cressie, 1985).

Os métodos de validação comparam valores teóricos do modelo geoestatístico escolhido e os valores empíricos obtidos na amostragem. Com base na análise desses erros de estimação, poderá ser selecionado o melhor modelo de semivariância. Entre os principais critérios para validação, encontram-se os de Informação de Akaike, de Filliben, de validação cruzada e o máximo valor do logaritmo da função verossimilhança.

O Critério de Informação de Akaike - AIC (*Akaike's Information Criterion*) procura uma solução satisfatória entre o bom ajuste e o princípio da parcimônia (Akaike, 1973). Sakamoto et al. (1986) desenvolveram estudos visando conhecer como os modelos são usados para fazer estimação. O AIC propõe verificar se dois modelos representam dados igualmente satisfatórios. Considera que, do modelo mais simples (menor valor de \hat{A}), pode-se esperar melhor desempenho para a estimação de novos dados, ou seja, o AIC impõe uma penalidade para a complexidade. O AIC é estimado por $\hat{A} = -2\ell + 2k$, em que ℓ é o logaritmo da função verossimilhança e k , o número de parâmetros do modelo ajustado. Para o caso de os dados apresentarem distribuição gaussiana de probabilidade, e se eliminadas as constantes arbitrárias, o AIC é estimado por $\hat{A} = n \ln \left(\frac{q_k}{n} \right) + 2k$, em que $q_k = \sum_{i=1}^n (Z(s_i) - \hat{Z}(s_i))^2$ é a soma de quadrados dos resíduos e n é o número de pontos amostrais.

O Critério de Filliben (1975) para resíduos ortonormais também pode ajudar na determinação do melhor modelo geoestatístico para o ajuste. Para a j -ésima posição ($j = p + 1, p + 2, \dots, n$), pode-se estimar por krigagem ordinária o valor de $\hat{Z}(s_j)$ usando somente os $j-1$ valores anteriores de dados e normalizados pelo desvio-padrão do erro da krigagem; o valor de p é obtido pela definição da função $\mu(s_j)$.

Os erros normalizados (Lee, 1994) para $n-p$ variáveis são definidos como:

$$\varepsilon_j = \frac{Z(s_j) - \hat{Z}(s_j)}{\sigma_j} \quad (2)$$

para $j = (p + 1), \dots, n$, em que σ_j é o desvio-padrão da estimativa da krigagem. Os $n-p$ resíduos obtidos pela equação 2 são chamados de resíduos ortonormais, isto é, são não-correlacionados, linearmente independentes e têm variância unitária. No teste de Filliben (1975), calcula-se o coeficiente de correlação linear r^* entre as observações ordenadas $\varepsilon_{(j)}$ e a estatística das medianas ordenadas $m_{(j)}$ de distribuição normal $N(0,1)$. Quanto mais próximo r^* estiver de 1, mais normal é a distribuição dos dados. Se o valor tabelado de r (Filliben, 1975) para determinado tamanho de amostra é maior do que r^* , a hipótese de que os dados seguem uma distribuição normal é rejeitada para um certo nível percentual de significância. Praticamente, o que se testa é se os resíduos ortonormais seguem uma

distribuição normal em determinado nível de significância.

A validação cruzada, segundo Isaaks & Srivastava (1989), é uma técnica de avaliação de erros de estimativas que permite comparar os valores previstos com os amostrados. O valor da amostra, em certa localização $Z(s_j)$, é temporariamente descartado do conjunto de dados e, então, é feita uma previsão por krigagem no local $\hat{Z}(s_{(j)})$, usando-se as amostras restantes. Assim, o Erro Médio por validação cruzada (EM) é obtido da equação 3:

$$EM = \frac{1}{n} \sum_{i=1}^n (Z(s_i) - \hat{Z}(s_{(i)})) \quad (3)$$

em que n é o número de dados; $Z(s_j)$, valor observado no ponto s_j ; e $\hat{Z}(s_{(j)})$, valor predito por krigagem ordinário no ponto s_j , sem considerar a observação $Z(s_j)$.

Esse procedimento pode ser visto como um experimento no qual se imita o processo de estimação, ao supor que nunca se toma uma amostra naquela localização. Uma vez que a estimação é feita, pode-se compará-la ao valor da amostra que foi inicialmente removida do conjunto de dados amostrais. Este procedimento, método de "deixar um fora", é repetido para todas as amostras disponíveis.

McBratney & Webster (1986) e Cressie (1993) apresentam erro médio reduzido (\overline{ER}), desvio-padrão dos erros médios (DP_{EM}), desvio-padrão dos erros reduzidos (S_{ER}) e do erro absoluto (EA), como instrumento para avaliar modelos. O erro médio reduzido (\overline{ER}) é definido pela equação 4:

$$\overline{ER} = \frac{1}{n} \sum_{i=1}^n \frac{Z(s_i) - \hat{Z}(s_{(i)})}{\sigma(\hat{Z}(s_{(i)}))} \quad (4)$$

em que $\sigma(\hat{Z}(s_{(j)}))$ é o desvio-padrão da krigagem no ponto s_j , sem considerar a observação $Z(s_j)$.

O desvio-padrão dos erros reduzidos (S_{ER}) é obtido a partir da Equação (5):

$$S_{ER} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{|Z(s_i) - \hat{Z}(s_{(i)})|}{\sigma(\hat{Z}(s_{(i)}))}} \quad (5)$$

Segundo McBratney & Webster (1986), Cressie (1993) e Mello et al. (2005), aplicando-se a condição de não-tendenciosidade, o valor populacional para o erro médio reduzido deve ser zero e o do desvio-padrão do erro reduzido igual a 1. Portanto, o valor de EM e \overline{ER} mais próximo de zero, o valor DP_{EM} menor e o valor de S_{ER} mais próximo de um são os critérios para escolha do melhor modelo ajustado. O erro absoluto (EA) é uma medida da magnitude dos erros na unidade da variável. Conhecendo-se o conjunto de valores medidos e preditos por krigagem ordinária $Z(s_j)$ e $\hat{Z}(s_{(j)})$, respectivamente, foi possível definir o erro absoluto na unidade da variável estudada, pela equação 6:

$$EA = \sum_{i=1}^n |Z(s_i) - \hat{Z}(s_i)| \quad (6)$$

Os métodos de estimação de mínimos quadrados ordinários (OLS) e mínimos quadrados ponderados (WLS1) buscam minimizar a soma de quadrados das diferenças entre valores observados e estimados. Já o método de estimação de máxima verossimilhança considerado um estimador de parâmetros assintoticamente normal, não viciado e eficiente (Cox & Hinkley, 1974) – consiste em maximizar a função de densidade de probabilidade conjunta do processo $Z(s)$ em relação aos efeitos fixos (processo estacionário) e os componentes dos efeitos aleatórios. McBratney & Webster (1986) utilizaram este método de ajuste de modelos e também afirmaram ser o mais eficiente. Mardia & Marshall (1984) desenvolveram teoria para o caso em que o termo aleatório $\epsilon(s)$ fosse um processo gaussiano. O melhor modelo para um processo será aquele que apresentar o maior valor de maximização do logaritmo da função verossimilhança.

O objetivo deste trabalho foi descrever os comportamentos espaciais dos dados de umidade do solo, densidade do solo e resistência do solo à penetração nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m e da produtividade da soja, pela seleção de modelos de variabilidade espacial, usando os métodos de estimação de mínimos quadrados ordinários (OLS), mínimos quadrados ponderados (WLS1) (Cressie, 1985) e máxima verossimilhança (MV) (Mardia & Marshall, 1984), segundo os critérios de Akaike, Filliben, validação cruzada e máximo valor do logaritmo da função verossimilhança (*MLL*). O trabalho também apresenta os mapas temáticos utilizando estrutura de dependência espacial, escolhida segundo os critérios utilizados.

MATERIAL E MÉTODOS

Os dados experimentais foram obtidos no ano agrícola 2004/2005, referentes a uma pesquisa que se desenvolveu numa propriedade localizada no município de Cascavel, Estado do Paraná, em área de produção de grãos de 57 ha, de um solo classificado como Latossolo Vermelho distroférrico. A variedade da soja semeada na área em estudo foi a COODETEC 216 (CD216). Na área experimental, cultivada sob plantio direto, foram demarcadas 100 parcelas com espaçamento de 75 x 75 m, com auxílio de um aparelho GPS, pelo método estático, com correção diferencial pós-processada, visando a sua correta localização no sistema de coordenadas geográficas Universal Transverse Mercatur (UTM), que utiliza coordenadas métricas. Em cada ponto foram: (a) coletadas amostras de solo para avaliação da umidade e densidade do solo nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m; (b) feitos furos com um penetrógrafo e registrada a resistência do solo à penetração nas camadas de 0 a

0,1, 0,1 a 0,2 e 0,2 a 0,3 m; e (c) colhida a soja da parcela e determinada a sua produtividade.

Para análise da estrutura de dependência espacial dos processos intrinsecamente estacionários e isotrópicos, utilizaram-se semivariogramas experimentais, construídos empregando-se o estimador de Matheron ou Cressie & Hawkins, caso os dados apresentem valores discrepantes, conforme o caso (Cressie, 1985). Com a finalidade de se aplicarem os critérios de validação em estudo, ajustaram-se três modelos teóricos ao semivariograma experimental: exponencial, esférico e gaussiano, considerados adequados aos dados em análise. Na estimação dos parâmetros foram usados os métodos: dos mínimos quadrados ordinários (OLS), dos mínimos quadrados ponderados (WLS1) (Cressie, 1985) e máxima verossimilhança (MV) (Mardia & Marshall, 1984). Para estimação e ajuste de modelos e avaliação de critérios de ajuste foi utilizado o *software R* (R Development Core Team, 2005) e, nele, o pacote *geoR* (Ribeiro Jr. & Diggle, 2001), ambos livres e de acordo com a licença GPL (General Public Licence).

Com a comparação entre os resultados alcançados, foram analisados os comportamentos dos critérios de Akaike, de Filliben, de validação cruzada e maior valor de maximização do logaritmo da função verossimilhança. Finalmente, construíram-se os mapas temáticos da umidade do solo, da densidade, da resistência do solo à penetração e da produtividade da soja, segundo o modelo que apresentou melhor ajuste utilizando-se os critérios de validação de modelos.

RESULTADOS E DISCUSSÃO

No quadro 1 são apresentadas as estatísticas descritivas da densidade do solo (DS), da resistência do solo a penetração (RSP) e da umidade do solo (UMD) nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m e da produtividade da soja (PROD). Para a densidade do solo, observa-se que os valores não apresentaram variações acentuadas, estando a média dentro de valores esperados para solos na região estudada: da ordem de 1,0 a 1,45 kg m⁻³. Os coeficientes de variação indicam homogeneidade (CV menor que 10 %), segundo Gomes & Garcia (2002). Para as três camadas estudadas, a média e mediana foram semelhantes, mostrando que as densidades do solo não variam muito entre as camadas. Para a RSP, observou-se, na camada de 0 a 0,1 m, estatísticas de posição superiores às das outras profundidades, notadamente a média e a mediana, porém com maior homogeneidade quando comparada às outras duas camadas que tiveram valores de coeficiente de variação entre 20 e 30 %. Observa-se ainda que os valores da RSP, na camada de 0 a 0,1 m, encontram-se entre 0,671 e 4,269 MPa. Os valores mais elevados ocorreram nas parcelas 10, 92, 96 e 97.

No estudo da umidade do solo (Quadro 1), a umidade média encontrada foi de $0,369 \text{ m}^3 \text{ m}^{-3}$, com desvio-padrão de $0,06 \text{ m}^3 \text{ m}^{-3}$ e coeficiente de variação de 15,56 % para a camada de 0 a 0,1 m. Para as outras duas camadas houve aumento médio da umidade, porém nas três camadas houve média homogeneidade (CV entre 10 e 20 %), segundo Gomes & Garcia (2002).

A produtividade média da soja (Quadro 1) foi de $3,22 \text{ t ha}^{-1}$, com desvio-padrão de 0,38. O valor mínimo encontrado foi de $2,09 \text{ t ha}^{-1}$, e o máximo, de $4,09 \text{ t ha}^{-1}$. Observa-se portanto que, em média, a produtividade para a área em estudo foi superior às médias no Estado

do Paraná e nacional, no ano agrícola 2004/2005, já que a produção média estadual, para essa safra, foi de $2,30 \text{ t ha}^{-1}$, e a média nacional, de $2,19 \text{ t ha}^{-1}$. O coeficiente de variação amostral (CV) encontrado para a produtividade da soja foi de 11,71 %, existindo, portanto, média homogeneidade nos dados em relação à sua média (Gomes & Garcia, 2002).

No quadro 2 apresentam-se os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) e alcance (a) dos modelos ajustados para os dados de densidade do solo nas camadas de 0 a 0,1 m, 0,1 a 0,2 m e 0,2 a 0,3 m, onde se observa que os dois primeiros parâmetros (efeito pepita e patamar) apresentam semelhança nos três

Quadro 1. Análise exploratória dos atributos densidade do solo (DS), resistência do solo à penetração (RSP) e umidade do solo (UMD) nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m e da produtividade da soja

Variável/ estatística	Densidade			Resistência à penetração			Produtividade	Umidade		
	DS ₀₋₁₀	DS ₁₀₋₂₀	DS ₂₀₋₃₀	RSP ₀₋₁₀	RSP ₁₀₋₂₀	RSP ₂₀₋₃₀		UMD ₀₋₁₀	UMD ₁₀₋₂₀	UMD ₂₀₋₃₀
	kg dm ⁻³			MPa			t ha ⁻¹	m ³ m ⁻³		
N	83	88	91	93	90	88	66	84	88	91
Média	1,101	1,118	1,077	2,729	2,427	2,101	3,217	0,369	0,391	0,406
Mínimo	0,770	0,980	0,950	0,671	0,875	0,770	2,090	0,230	0,220	0,210
1° Quartil	1,050	1,050	1,030	2,374	2,131	1,767	2,960	0,348	0,360	0,390
Mediana	1,100	1,110	1,070	2,692	2,330	2,031	3,190	0,375	0,400	0,420
3° Quartil	1,160	1,160	1,120	3,125	2,771	2,291	3,480	0,403	0,423	0,440
Máximo	1,310	1,550	1,300	4,269	3,709	4,958	4,090	0,590	0,600	0,640
DP	0,080	0,090	0,070	0,670	0,510	0,530	0,380	0,060	0,060	0,070
CV (%)	7,590	7,690	6,420	24,50	21,04	25,38	11,71	15,56	16,17	16,38
CVT (%)	7,270	8,110	6,540	24,89	21,89	26,10	11,91	16,00	15,00	16,67
Assimetria	-0,046	1,686	0,737	-0,170	0,111	2,040	-0,180	-0,031	0,038	-0,670
Curtose	2,030	6,225	0,790	0,692	0,444	9,310	0,696	2,213	2,406	2,204

N: número de amostras; DP: desvio-padrão; CV: coeficiente de variação; CVT: coeficiente de variação de Thorndike (100 DP/Mediana); as diferenças para os valores no número de pontos N ocorreram devido à perda de informação.

Quadro 2. Parâmetros dos modelos semivariográficos, pelos métodos mínimos quadrados ordinários (OLS), mínimos quadrados ponderados (WLS1) e máxima verossimilhança (MV) para densidade do solo (DS) nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m

Camada	Modelo	Efeito pepita (C_0) x 10 ⁻³			Patamar ($C_0 + C_1$) x 10 ⁻³			Alcance (a)		
		OLS	WLS1	MV	OLS	WLS1	MV	OLS	WLS1	MV
		(kg dm ⁻³) ²						m		
0 a 0,1 m	Exponencial	5,0	6,3	0,0	7,7	7,7	6,9	95,0	120,0	50,5
	Esférico	6,0	5,5	5,8	7,7	7,7	6,9	126,0	140,0	106,8
	Gaussiano	3,5	5,0	1,3	7,7	7,7	6,8	87,0	97,0	115,0
	Exponencial	5,0	5,0	7,2	8,3	8,0	7,4	239,0	120,0	110,2
0,1 a 0,2 m	Esférico	6,0	5,0	6,7	8,0	8,0	7,3	150,0	120,0	90,0
	Gaussiano	5,0	5,0	6,9	7,9	8,0	7,3	73,0	78,0	135,0
	Exponencial	3,5	3,0	4,7	4,5	4,5	4,8	50,0	141,0	16,8
0,2 a 0,3 m	Esférico	3,5	3,0	4,2	4,5	4,5	4,7	78,0	140,0	75,1
	Gaussiano	3,0	2,5	2,8	4,5	4,5	3,3	75,0	139,0	32,3

métodos de estimação (OLS, WLS1 e MV) para os três modelos em estudo (exponencial, esférico e gaussiano). Utilizando o método da máxima verossimilhança (MV), observa-se que os parâmetros estimados nas camadas de 0,1 a 0,2 e 0,2 a 0,3 m indicam que os modelos têm efeito pepita puro, isto é, há pouca diferença entre o efeito pepita (C_0) e o patamar ($C_0 + C_1$).

No quadro 3, encontram-se os resultados dos diferentes critérios de validação de ajustes dos modelos geoestatísticos. Segundo o critério AIC, o menor valor de \hat{A} para a DS na camada de 0 a 0,1 m aponta para o modelo gaussiano estimado pelos métodos OLS e WLS1. Para a DS na camada de 0,1 a 0,2 m, o menor valor de \hat{A} indica o modelo exponencial estimado pelos métodos OLS e WLS1; para a DS na camada de 0,2 a 0,3 m, o menor valor de \hat{A} aponta para os três modelos estimados pelo método OLS e esférico para WLS1. O critério de Filliben testou se os resíduos ortonormais seguiam uma distribuição normal a 1 %. Para o caso dos dados da DS nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m, que têm 83, 88 e 91 valores amostrados, respectivamente, os valores tabelados de r , obtidos da tabela de Filliben (1975), foram de 0,976, 0,977 e 0,978, respectivamente. Comparando com os valores r^* calculados (Quadro 3), verificou-se que os valores r^* são maiores do que r tabelado, indicando que a hipótese de que os dados seguem distribuição normal não foi rejeitada para um nível de 1 % de significância. Pelo critério de Filliben, para a DS nas camadas de 0 a 0,1 e 0,2 a 0,3 m todos os modelos são aceitos, e para DS na camada de 0,1 a 0,2 m, nenhum modelo foi aceito. Pelo máximo valor do logaritmo da função verossimilhança (MLL), os valores são semelhantes em cada camada, não se destacando nenhum modelo.

O critério de validação cruzada utilizando o erro médio da validação cruzada (EM), erro médio reduzido (\overline{ER}), desvio-padrão dos erros médios (DP_{EM}), desvio-

padrão dos erros reduzidos (S_{ER}) e do erro absoluto (EA), apresentados no quadro 4, aplicados aos modelos em estudo, indicou, para as três camadas os modelos esférico, exponencial e exponencial, com parâmetros estimados pelos métodos WLS1, MV e OLS, respectivamente.

A figura 1 apresenta os mapas temáticos da densidade do solo (DS) nas três camadas, construídos utilizando os modelos indicados pelo critério da validação cruzada. Observa-se nesta figura (a,b,c) que, à medida que a profundidade aumenta, há diminuição na densidade do solo em $[\text{kg dm}^{-3}]$.

Os semivariogramas experimentais para a variável resistência do solo à penetração (RSP), nas camadas de 0 a 0,1 e 0,1 a 0,2 m, foram calculados pelo estimador de Matheron, e para a camada de 0,2 a 0,3 m, pelo estimador de Cressie & Hawkins, por apresentarem valores discrepantes. No quadro 5 são apresentados os parâmetros efeito pepita (C_0), patamar ($C_0 + C_1$) e alcance (a) para os modelos ajustados para a RSP. Observou-se que esses parâmetros apresentam semelhança nos três métodos de estimação (OLS, WLS1 e MV) nos três modelos ajustados.

No quadro 6 apresenta-se avaliação dos modelos da RSP. Os valores \hat{A} do critério AIC apontaram, para a camada de 0 a 0,1 m, o modelo esférico utilizando os três métodos de estimação (OLS, WLS1 e MV); para a camada de 0,1 a 0,2 m, escolheu-se o modelo esférico com os três métodos de estimação e também o modelo gaussiano com o método de MV; para a camada de 0,2 a 0,3 m, escolheu-se o modelo esférico com o método de MV e gaussiano com os métodos OLS e WLS1. Utilizando o critério de Filliben, para a RSP nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m, com 93, 90 e 88 valores amostrados, respectivamente, e para um nível de significância de 1 %, os valores de r obtidos da tabela apresentada em

Quadro 3. Resultados dos critérios de AIC (\hat{A}), Filliben (r^*) e valor máximo do logaritmo da função verossimilhança (MLL) para a densidade do solo (DS) nas camadas de 0 a 0,1; 0,1 a 0,2; e 0,2 a 0,3 m

Camada	Modelo	N	\hat{A}			r^*			MLL
			OLS	WLS1	MV	OLS	WLS1	MV	
DS 0 a 0,10 m	Exponencial	83	-1109,8	-640,3	-	<u>0,9780</u>	<u>0,9761</u>	<u>0,9787</u>	88,80
	Esférico	83	-1115,2	-651,2	-	<u>0,9773</u>	<u>0,9760</u>	<u>0,9786</u>	88,80
	Gaussiano	83	<u>-1117,0</u>	<u>-652,8</u>	-	<u>0,9766</u>	<u>0,9767</u>	<u>0,9786</u>	88,80
DS 0,1 a 0,2 m	Exponencial	88	<u>-1136,5</u>	<u>-641,7</u>	-	0,9439	0,9409	0,9391	91,57
	Esférico	88	-1119,8	-632,9	-	0,9420	0,9413	0,9392	91,57
	Gaussiano	88	-1118,8	-635,3	-	0,9400	0,9399	0,9392	91,57
DS 0,2 a 0,3 m	Exponencial	91	<u>-1217,8</u>	-717,9	-	<u>0,9818</u>	<u>0,9851</u>	<u>0,9818</u>	114,40
	Esférico	91	<u>-1217,8</u>	<u>-720,2</u>	-	<u>0,9818</u>	<u>0,9856</u>	<u>0,9818</u>	114,40
	Gaussiano	91	<u>-1217,8</u>	-700,0	-	<u>0,9818</u>	<u>0,9890</u>	<u>0,9816</u>	117,50

N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; MV: máxima verossimilhança; r^* : teste de Filliben compara-se com $r = 0,976$ (0 a 0,1 m), $0,977$ (0,1 a 0,2 m) e $0,978$ (0,2 a 0,3 m) a 1 %.

Quadro 4. Validação cruzada para a densidade do solo (DS) [kg dm⁻³] nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m

Variável	Modelo	Método	EM	\overline{ER}	DP _{EM}	S _{ER}	EA
DS 0 a 0,1 m	Exponencial	OLS	-0,0000040	-0,0000225	0,08461952	0,96455996	5,447215
		WLS1	-0,0000033	-0,0000186	0,08455092	0,96171235	5,445539
		MV	-0,0000098	-0,0000059	0,08448363	1,01187008	5,442573
	Esférico	OLS	-0,0000070	-0,0000395	0,0845969	0,9688203	5,496190
		WLS1	<u>-0,0000002</u>	<u>-0,0000016</u>	<u>0,08452754</u>	<u>0,96079735</u>	<u>5,442286</u>
		MV	-0,0000015	-0,0000088	0,08444795	1,01181780	5,448480
	Gaussiano	OLS	-0,0000092	-0,0000517	0,08465853	0,97160520	5,502601
		WLS1	-0,0000091	-0,0000516	0,08450482	0,96502022	5,454448
		MV	-0,0000014	-0,0000087	0,08445353	1,01182725	5,447732
DS 0,1 a 0,2 m	Exponencial	OLS	-0,0000792	-0,0004433	0,08822456	1,00106809	5,650872
		WLS1	-0,0001338	-0,0007592	0,0888563	1,0205829	5,720422
		MV	<u>-0,0000005</u>	<u>-0,0000031</u>	<u>0,08693769</u>	<u>1,01140552</u>	<u>5,551424</u>
	Esférico	OLS	-0,0000070	-0,0000392	0,0869169	0,9691241	5,590005
		WLS1	-0,0000106	-0,0000593	0,0869194	0,9681909	5,585087
		MV	-0,0000009	-0,0000050	0,08694189	1,01142987	5,609558
	Gaussiano	OLS	-0,0000512	-0,0002863	0,08754412	0,98255020	5,594473
		WLS1	-0,0000423	-0,0002363	0,08724685	0,97629707	5,567138
		MV	-0,0000008	-0,0000048	0,0869413	1,0114267	5,609322
DS 0,2 a 0,3 m	Exponencial	OLS	<u>0,00000002</u>	<u>0,00000012</u>	<u>0,0699668</u>	<u>1,0461129</u>	<u>4,921846</u>
		WLS1	0,00001505	0,0001136	0,07094138	1,06865748	5,006869
		MV	0,00000000	0,0000000	0,0699641	1,0111111	4,921556
	Esférico	OLS	0,00000194	0,00001450	0,07014228	1,04989764	4,935683
		WLS1	0,000008669	0,00006556	0,07054922	1,06439371	4,959607
		MV	0,00000000	0,00000000	0,0699641	1,0111111	4,921555
	Gaussiano	OLS	0,000000107	0,00000080	0,06998064	1,04628826	4,923208
		WLS1	0,00001324	0,0001006	0,07084687	1,07329040	4,981735
		MV	0,00009282	0,00281900	0,0699414	1,0500280	4,921247

EM: erro médio da validação cruzada; \overline{ER} : erro médio reduzido; DP_{EM}: desvio-padrão dos erros médios; S_{ER}: desvio-padrão dos erros reduzidos; EA: erro absoluto.

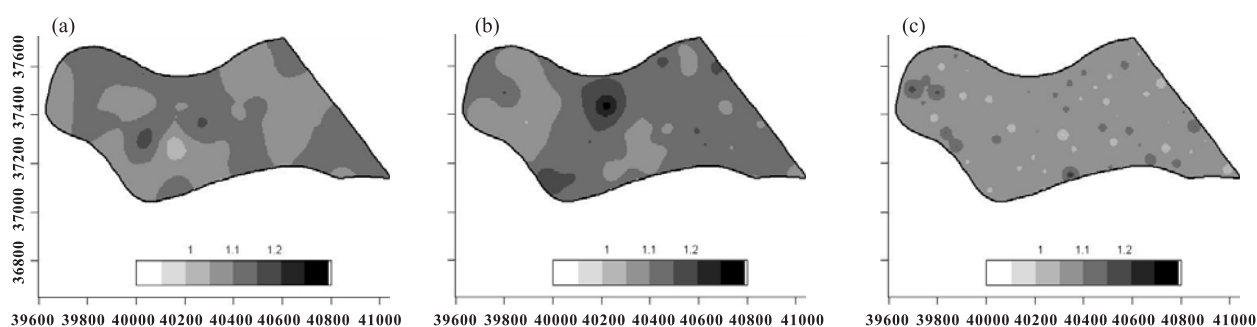


Figura 1. Mapa temático para a densidade do solo [kg dm⁻³] nas camadas de 0 a 0,1 m (a), 0,1 a 0,2 m (b) e 0,2 a 0,3 m (c).

Filliben (1975) foram de 0,978, 0,978 e 0,978, nessa ordem. Comparando com os valores de r^* apresentados no quadro 6, verificou-se que estes são maiores do que r nas duas primeiras camadas, indicando que os erros seguem distribuição normal em um nível de significância de 1%. Por esse critério, nas camadas

de 0 a 0,1 e 0,1 a 0,2 m, todos os modelos são aceitos, mas, para a camada de 0,2 a 0,3 m, nenhum modelo foi aprovado. Pelo máximo valor do logaritmo da função verossimilhança (*MLL*), observa-se que os valores são semelhantes em cada camada, não se destacando nenhum modelo.

Quadro 5. Parâmetros dos modelos semivariográficos pelos métodos mínimos quadrados ordinários (OLS), mínimos quadrados ponderados (WLS1) e máxima verossimilhança (MV) para os dados de RSP nas camadas de 0 a 0,1; 0,1 a 0,2; e 0,2 a 0,3 m

Camada	Modelo	Efeito pepita (C_0)			Patamar (C_0+C_1)			Alcance (a)		
		OLS	WLS1	MV	OLS	WLS1	MV	OLS	WLS1	MV
		(MPa) ²						m		
0 a 0,1 m	Exponencial	0,3594	0,3518	0,0000	0,4464	0,4399	0,4436	638	450	109
	Esférico	0,3649	0,3700	0,3914	0,4382	0,4378	0,4464	445	440	1172
	Gaussiano	0,3800	0,3800	0,3176	0,4400	0,4375	0,4448	426	369	262
0,1 a 0,2 m	Exponencial	0,1894	0,1699	0,2052	0,2870	0,2782	0,2572	901	600	363
	Esférico	0,2006	0,1969	0,2509	0,2765	0,2752	0,2577	647	600	166
	Gaussiano	0,2144	0,2123	0,2219	0,2778	0,2771	0,2582	581	554	568
0,2 a 0,3 m	Exponencial	0,0000	0,0000	0,0000	0,2580	0,2607	0,2815	784	824	110
	Esférico	0,0000	0,0000	0,2426	0,2397	0,2401	0,2844	591	609	1383
	Gaussiano	0,1184	0,1175	0,2523	0,2749	0,2864	0,2836	901	946	708

Quadro 6. Resultados dos critérios de AIC (\hat{A}), Filliben (r^*) e máximo valor do logaritmo da função verossimilhança (MLL) para a RSP nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m

Variável	Modelo	N	\hat{A}			r^*			MLL
			OLS	WLS1	MV	OLS	WLS1	MV	
RSP 0,0 a 0,1 m	Exponencial	93	-416,8	108,4	193,9	<u>0,9885</u>	<u>0,9885</u>	<u>0,9895</u>	-92,95
	Esférico	93	<u>-420,7</u>	<u>103,9</u>	<u>193,3</u>	<u>0,9889</u>	<u>0,9889</u>	<u>0,9892</u>	-93,39
	Gaussiano	93	-419,3	105,6	193,7	<u>0,9890</u>	<u>0,9890</u>	<u>0,9893</u>	-92,83
RSP 0,10 a 0,2 m	Exponencial	90	-589,3	-88,4	138,6	<u>0,9859</u>	<u>0,9871</u>	<u>0,9851</u>	-65,31
	Esférico	90	<u>-599,5</u>	<u>-105,6</u>	<u>138,2</u>	<u>0,9849</u>	<u>0,9854</u>	<u>0,9851</u>	-66,69
	Gaussiano	90	-595,6	-103,6	<u>138,2</u>	<u>0,9853</u>	<u>0,9857</u>	<u>0,9857</u>	-65,08
RSP 0,20 a 0,3 m	Exponencial	88	-473,4	19,8	143,8	0,9200	0,9201	0,9080	-67,90
	Esférico	88	-474,3	19,5	<u>142,6</u>	0,9194	0,9198	0,9095	-67,92
	Gaussiano	88	<u>-490,2</u>	<u>1,1</u>	144,0	0,9095	0,9096	0,9017	-68,02

N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; MV: máxima verossimilhança; r^* , teste de Filliben, compara-se com $r = 0,978$ (0 a 0,1 m), $0,978$ (0,1 a 0,2 m) e $0,978$ (0,2 a 0,3 m) a 1%.

O critério de validação cruzada apresentado no quadro 7 sugeriu para a RSP a escolha dos modelos exponencial, esférico e gaussiano, respectivamente, para as três camadas. Nas três escolhas, o método de estimação de parâmetros foi o MV.

Na figura 2 são apresentados os mapas temáticos da resistência do solo à penetração nas três camadas em estudo, construídos com os modelos indicados pelo critério da validação cruzada. Observa-se que, à medida que a profundidade aumenta, há diminuição da RSP em [MPa] e homogeneização espacial da área.

No quadro 8 encontram-se os modelos ajustados e os parâmetros estimados efeito pepita (C_0), patamar ($C_0 + C_1$) e alcance (a) para a umidade do solo (UMD) nas três camadas em estudo. A estimação do semivariograma experimental para as camadas de 0

a 0,1 e 0,2 a 0,3 m foi calculada pelo estimador de Cressie & Hawkins, pois os dados mostraram pontos discrepantes; já na camada de 0,1 a 0,2 m, o semivariograma experimental foi calculado pelo estimador de Matheron. Observa-se também que os dois primeiros parâmetros (efeito pepita e patamar) apresentam semelhança nos três métodos de estimação (OLS, WLS1 e MV) para os três modelos ajustados.

No quadro 9 é apresentada a avaliação dos modelos da UMD. Os valores \hat{A} do critério AIC apontaram para a camada de 0 a 0,1 m os modelos exponencial, empregando método de estimação WLS1, e gaussiano, utilizando os métodos OLS e MV; para a camada de 0,1 a 0,2 m, os modelos exponencial, com método de estimação MV, e gaussiano, utilizando os métodos OLS e WLS1; e para a camada de 0,2 a 0,3 m, os modelos

Quadro 7. Validação cruzada para a resistência do solo à penetração (RSP) [MPa] nas camadas de 0 a 0,1; 0,1 a 0,2; e 0,2 a 0,3 m

Variável	Modelo	Mét	EM	\overline{ER}	DP _{EM}	S _{ER}	EA
RSP 0, a 0,1m	Exponencial	OLS	0,00074	0,00058	0,66317	1,02956	46,14848
		WLS	0,00074	0,00058	0,66338	1,02460	46,25310
		MV	<u>0,00052</u>	<u>0,00040</u>	<u>0,65826</u>	<u>1,00912</u>	46,22873
	Esférico	OLS	0,00052	0,00041	0,65712	1,02041	46,30269
		WLS	0,00053	0,00041	0,65711	1,02016	46,29807
		MV	0,00064	0,00050	0,65311	1,00929	45,55866
	Gaussiano	OLS	0,00065	0,00051	0,66244	1,03017	46,06193
		WLS	0,00064	0,00050	0,66251	1,02616	46,14322
		MV	0,00056	0,00043	0,65608	1,00919	45,93212
RSP 0,1 a 0,2 m	Exponencial	OLS	-0,00035	-0,00036	0,50422	1,04388	35,08883
		WLS	-0,00043	-0,00044	0,50526	1,05685	35,16094
		MV	-0,00007	0,00007	0,50123	1,00890	35,03478
	Esférico	OLS	-0,00073	-0,00074	0,51342	1,03655	36,42900
		WLS	-0,00010	-0,00009	0,50406	1,04515	34,63800
		MV	<u>0,00002</u>	<u>0,00002</u>	<u>0,49916</u>	<u>1,00917</u>	34,89847
	Gaussiano	OLS	-0,00031	-0,00032	0,50412	1,04332	35,44404
		WLS	0,00002	0,00004	0,50351	1,04291	34,44807
		MV	0,00005	0,00005	0,49890	1,00909	34,83145
RSP 0,2 a 0,3 m	Exponencial	OLS	-0,00163	-0,00166	0,53183	1,11307	31,04510
		WLS	-0,00099	-0,00102	0,53042	1,11783	30,51272
		MV	-0,00035	-0,00034	0,52394	1,00827	30,25766
	Esférico	OLS	-0,00088	-0,00090	0,53013	1,10150	30,46770
		WLS	-0,00098	-0,00101	0,53041	1,11742	30,50507
		MV	-0,00029	-0,00026	0,51230	1,00718	29,21713
	Gaussiano	OLS	-0,00136	-0,00136	0,53627	1,09944	31,41471
		WLS	-0,00050	-0,00050	0,53014	1,11368	30,56658
		MV	<u>-0,00017</u>	<u>-0,00016</u>	<u>0,52673</u>	<u>1,01029</u>	30,56865

EM: erro médio da validação cruzada; \overline{ER} : erro médio reduzido; DP_{EM}: desvio-padrão dos erros médios; S_{ER}: desvio-padrão dos erros reduzidos; EA: erro absoluto.

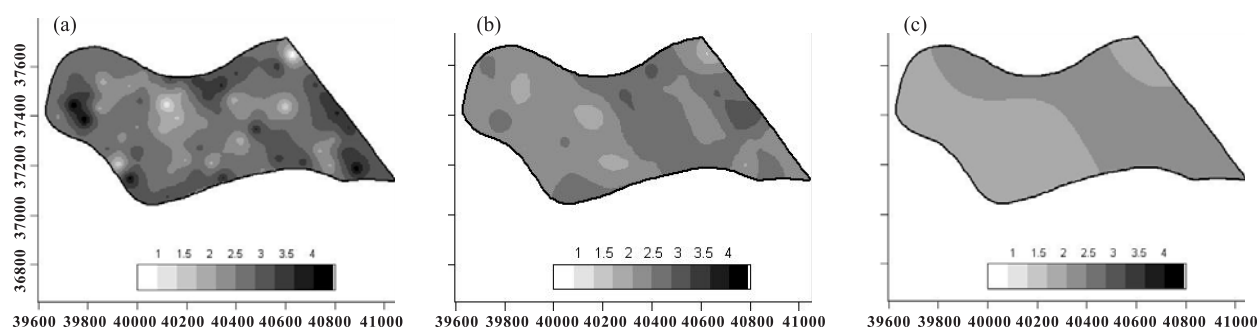


Figura 2. Mapa temático para a resistência do solo à penetração [MPa] nas camadas de 0 a 0,1 m (a), 0,1 a 0,2 m (b) e 0,2 a 0,3 m (c).

exponencial, utilizando método de estimação OLS, e esférico, empregando os métodos WLS1 e MV. Utilizando o critério de Filliben, para a camada de 0 a 0,1 m, a 1 %, foi escolhido o modelo gaussiano utilizando o estimador MV; na camada de 0,1 a 0,2 m,

o modelo esférico com o estimador WLS1; e na camada de 0,2 a 0,3 m, nenhum modelo foi aceito. Pelo máximo valor do logaritmo da função verossimilhança (MLL), observa-se que os valores são semelhantes em cada camada, não se destacando nenhum modelo.

O critério de validação cruzada apresentado no quadro 10 sugeriu, para a umidade na camada de 0 a 0,1 m, o modelo exponencial utilizando o método de estimação de MV; na camada de 0,1 a 0,2 m, para o modelo exponencial empregando o método de estimação de OLS; e na camada de 0,2 a 0,3 m, para o modelo esférico utilizando o método de estimação de MV.

Na figura 3 são apresentados os mapas temáticos referentes à UMD nas três camadas em estudo, construídos com base nos modelos indicados pelo critério da validação cruzada.

No quadro 11 apresentam-se os modelos ajustados e os parâmetros estimados efeito pepita (C_0), patamar ($C_0 + C_1$) e alcance (a), para a produtividade da soja [$t\ ha^{-1}$]. O semivariograma experimental foi calculado

pelo estimador de Matheron. Observa-se que os dois primeiros parâmetros (efeito pepita e patamar) apresentam semelhança nos três métodos de estimação (OLS, WLS1 e MV) para os três modelos ajustados.

No quadro 12, observa-se que, para a produtividade da soja, o critério AIC definiu o modelo gaussiano, com parâmetros estimados pelos métodos OLS e MV. Pelo critério de Filliben, com 66 valores amostrados, e para um nível de 1 %, o valor de r foi de 0,972 (Filliben, 1975). Por este critério, todos os modelos ajustados podem ser aceitos.

No quadro 13 é apresentado o critério de validação cruzada. Segundo este critério, escolheu-se o modelo exponencial com parâmetros estimados pelo método MV.

Quadro 8. Parâmetros dos modelos semivariográficos pelos métodos mínimos quadrados ordinários (OLS), mínimos quadrados ponderados (WLS1) e máxima verossimilhança (MV) para a umidade nas camadas de 0 a 0,1; 0,1 a 0,2; e 0,2 a 0,30 m

Camada	Modelo	Efeito pepita (C_0)			Patamar ($C_0 + C_1$)			Alcance (a)		
		OLS	WLS1	MV	OLS	WLS1	MV	OLS	WLS1	MV
		$(m^3\ m^{-3})^2$						m		
0 a 0,1 m	Exponencial	0,0000	0,0000	0,0026	0,0031	0,0031	0,0032	233	240	252
	Esférico	0,0009	0,0000	0,0027	0,0030	0,0031	0,0031	234	250	231
	Gaussiano	0,0013	0,0004	0,0028	0,0030	0,0032	0,0032	208	242	200
0,1 a 0,2 m	Exponencial	0,0023	0,0000	0,0000	0,0047	0,0048	0,0036	512	360	121
	Esférico	0,0026	0,0000	0,0014	0,0045	0,0048	0,0037	406	350	138
	Gaussiano	0,0030	0,0029	0,0012	0,0046	0,0046	0,0037	372	346	110
0,2 a 0,3 m	Exponencial	0,0010	0,0010	0,0038	0,0028	0,0029	0,0045	204	390	218
	Esférico	0,0015	0,0015	0,0040	0,0028	0,0029	0,0046	314	400	213
	Gaussiano	0,0024	0,0023	0,0042	0,0031	0,0028	0,0045	674	398	300

Quadro 9. Resultados dos critérios de AIC (\hat{A}), Filliben (r^*) e máximo valor do logaritmo da função verossimilhança (MLL) para a UMD nas camadas de 0 a 0,1; 0,1 a 0,2; e 0,2 a 0,3 m

Camada	Modelo	N	\hat{A}			r^*			MLL
			OLS	WLS1	MV	OLS	WLS1	MV	
UMD 0 a 0,1 m	Exponencial	84	-5981,2	<u>-5879,4</u>	-6282,2	0,95088	0,95031	0,94251	123,20
	Esférico	84	-5787,6	-5838,9	-5995,8	0,96664	0,96819	0,94183	123,10
	Gaussiano	84	<u>-6097,7</u>	-5821,8	<u>-6398,7</u>	0,94687	<u>0,98004</u>	0,94305	123,10
UMD 0,1 a 0,2 m	Exponencial	88	-6137,1	-6022,3	<u>-6361,2</u>	0,95455	0,96128	0,94920	123,90
	Esférico	88	-6111,0	-6195,0	-6058,9	0,97643	<u>0,98060</u>	0,95076	124,00
	Gaussiano	88	<u>-6176,4</u>	<u>-6281,6</u>	-6130,2	0,95800	0,95813	0,95033	124,00
UMD 0,2 a 0,3 m	Exponencial	91	<u>-6465,6</u>	-6378,3	---	0,92975	0,92842	0,91176	116,90
	Esférico	91	-6280,2	<u>-6606,3</u>	<u>-6639,5</u>	0,95742	0,92431	0,91216	116,90
	Gaussiano	91	-6416,8	-6439,5	---	0,92866	0,92743	0,91148	116,90

N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; MV: máxima verossimilhança; r^* : teste de Filliben, compara-se com $r = 0,977$ (0 a 0,1 m), $0,978$ (0,1 a 0,2 m) e $0,978$ (0,2 a 0,3 m) a 1 % de significância.

Quadro 10. Validação cruzada para a umidade do solo nas camadas de 0 a 0,1, 0,1 a 0,2 e 0,2 a 0,3 m

Variável	Modelo	Mét	EM	\overline{ER}	DP _{EM}	S _{ER}	EA
UMD 0,0 a 0,1 m	Exponencial	OLS	-0,0000345	-0,0003626	0,0563686	1,1801752	3,4897430
		WLS	-0,0000312	-0,0003265	0,0563304	1,1738025	3,4843900
		MV	<u>-0,0000171</u>	<u>-0,0001529</u>	<u>0,0563018</u>	<u>1,0093947</u>	3,4552240
	Esférico	OLS	-0,0000066	-0,0001494	0,0609737	1,4944216	3,779260
		WLS	-0,0000159	0,0000980	0,0616372	1,5586277	3,8331210
		MV	-0,0000211	-0,0001890	0,0563644	1,0099785	3,4423520
	Gaussiano	OLS	-0,0000217	-0,0001937	0,0564017	1,0098954	3,4858010
		WLS	0,0002003	0,0018650	0,0682667	1,8178146	4,3773130
		MV	-0,0000033	-0,0000292	0,0563314	1,0089312	3,4712120
UMD 0,1 a 0,2 m	Exponencial	OLS	<u>-0,0000248</u>	<u>-0,0002477</u>	0,0595221	1,1756804	3,6496470
		WLS	0,0002453	0,0028140	0,0607675	1,4455842	3,8642290
		MV	0,0000415	0,0003536	0,0589081	1,0081946	3,6350370
	Esférico	OLS	0,0004958	0,0062140	0,0657181	1,7438264	4,2639510
		WLS	0,0005938	0,0077700	0,0668942	1,8514137	4,3859960
		MV	0,0000871	0,0007452	0,0586952	1,0094746	3,6401430
	Gaussiano	OLS	-0,0000414	-0,0004208	0,0599438	1,2008590	3,6494750
		WLS	-0,0000393	-0,0004009	0,0599643	1,2020864	3,6494750
		MV	0,0000908	0,0007770	0,0587151	1,0094733	3,6414020
UMD 0,2 a 0,3 m	Exponencial	OLS	-0,0000865	-0,0009064	0,0686029	1,4392196	4,3157230
		WLS	-0,0000861	-0,0006938	0,0682887	1,0989296	4,2900370
		MV	-0,0000327	-0,0002437	0,0677356	1,0101709	4,2530580
	Esférico	OLS	-0,0002727	-0,0034550	0,0733873	1,9161152	4,7758340
		WLS	-0,0000677	-0,0005413	0,0682820	1,0891296	4,3026440
		MV	<u>-0,0000292</u>	<u>-0,0002174</u>	<u>0,0677459</u>	<u>1,0105066</u>	4,2550780
	Gaussiano	OLS	-0,0000827	-0,0008786	0,0681084	1,4354952	4,2881520
		WLS	-0,0000881	-0,0007134	0,0679966	1,0946607	4,2814970
		MV	-0,0000333	-0,0002490	0,0677243	1,0101827	4,2546860

EM: erro médio da validação cruzada; \overline{ER} : erro médio reduzido; DP_{EM}: desvio-padrão dos erros médios; S_{ER}: desvio-padrão dos erros reduzidos; EA: erro absoluto.

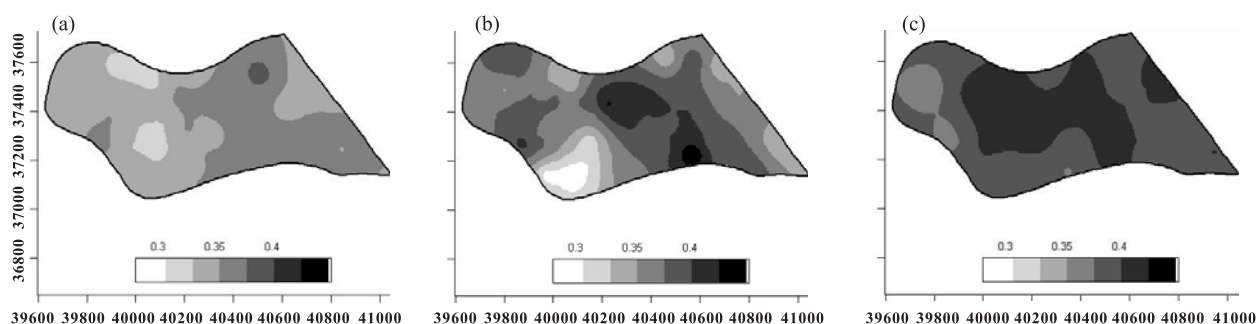


Figura 3. Mapa temático para a umidade do solo [m³ m⁻³] nas camadas de 0 a 0,1 (a), 0,1 a 0,2 (b) e 0,2 a 0,3 m (c).

A figura 4 apresenta o mapa temático de variabilidade espacial da produtividade da soja, construído com base no modelo indicado pelo critério da validação cru-

zada. Espacialmente, verifica-se que há regiões da área em estudo com produtividade superior à das médias estadual e nacional no ano agrícola 2004-2005.

Quadro 11. Parâmetros dos modelos semivariográficos pelos métodos mínimos quadrados ordinários (OLS), mínimos quadrados ponderados (WLS1) e máxima verossimilhança (MV) para a produtividade da soja

Modelo	Efeito pepita (C_0)			Patamar ($C_0 + C_1$)			Alcance (a)		
	OLS	WLS1	MV	OLS	WLS1	MV	OLS	WLS1	MV
	$(t\ ha^{-1})^2$						m		
Exponencial	0,0096	0,0000	0,0424	0,1519	0,1532	0,1361	341	320	262
Esférico	0,0394	0,0358	0,0747	0,1477	0,1490	0,1353	309	300	877
Gaussiano	0,0624	0,0664	0,0871	0,1485	0,1505	0,1359	283	294	461

Quadro 12. Resultados dos critérios de AIC (\hat{A}), Filliben (r^*) e máximo valor do logaritmo da função verossimilhança (MLL) para a produtividade da soja

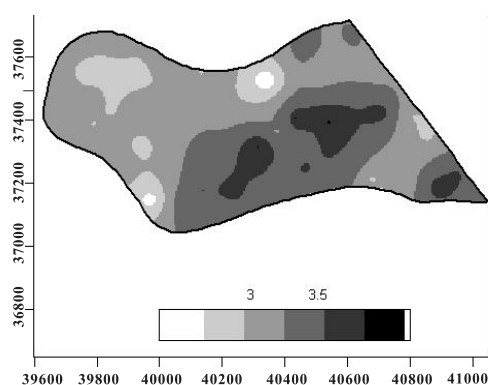
Variável	Modelo	N	\hat{A}			r^*			MLL
			OLS	WLS1	MV	OLS	WLS1	MV	
Produtividade (Mg ha ⁻¹)	Exponencial	66	-402,1	-77,1	56,1	<u>0,9866</u>	<u>0,9879</u>	<u>0,9793</u>	-24,04
	Esférico	66	-403,9	-76,2	57,6	<u>0,9837</u>	<u>0,9848</u>	<u>0,9853</u>	-23,90
	Gaussiano	66	<u>-405,5</u>	-76,9	<u>55,8</u>	<u>0,9837</u>	<u>0,9823</u>	<u>0,9777</u>	-23,90

N: número de pontos amostrados; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; MV: máxima verossimilhança; r^* : teste de Filliben compara-se com $r = 0,972$ a 1%.

Quadro 13. Validação cruzada para produtividade

Variável	Modelo	Método	EM	\overline{ER}	DP _{EM}	S _{ER}	EA
Produtividade Mg ha ⁻¹	Exponencial	OLS	-0.00277	-0.00400	0.34925	1.06195	16.66196
		WLS	-0.00225	-0.00318	0.34788	1.02730	16.50902
		MV	<u>-0.00173</u>	<u>-0.00242</u>	<u>0.34473</u>	<u>1.01735</u>	16.51318
	Esférico	OLS	-0.00837	-0.01298	0.37258	1.33974	18.79747
		WLS	-0.00872	-0.01382	0.37361	1.37521	18.71498
		MV	-0.00175	-0.00248	0.34386	1.01702	16.44871
	Gaussiano	OLS	-0.00307	-0.00453	0.34879	1.08066	16.77925
		WLS	-0.00250	-0.00359	0.34698	1.04065	16.62375
		MV	-0.00186	-0.00263	0.34379	1.01712	16.44577

EM: erro médio da validação cruzada; \overline{ER} : erro médio reduzido; DP_{EM}: desvio-padrão dos erros médios; S_{ER}: desvio-padrão dos erros reduzidos; EA: erro absoluto.

**Figura 4. Mapa temático para a produtividade da soja [t ha⁻¹].**

Finalmente, segundo o resumo do quadro 14, o critério da validação cruzada foi considerado o mais seletivo; já o critério de Filliben foi o mais conservador, e o critério de Akaike foi pouco conclusivo.

CONCLUSÕES

1. Os critérios de seleção por validação cruzada, Akaike, Filliben e máximo valor do logaritmo da função verossimilhança não selecionam simultaneamente um mesmo modelo, devido em parte às distintas naturezas dos critérios. A aplicação do critério de Filliben não apontou para um modelo particular em nenhum

Quadro 14. Resumos dos modelos escolhidos com melhor ajuste

Atributo	Critério de Akaike	Critério de Filliben	Validação cruzada
DS 0 a 0,1 m	GAU (OLS, WLS1)	TODOS (OLS, WLS1, MV)	ESF (WLS1)
DS 0,1 a 0,2 m	EXP (OLS, WLS1)	NENHUM	EXP (MV)
DS 0,2 a 0,3 m	EXP (OLS); GAU (OLS) ESF (OLS,WLS1)	TODOS (OLS, WLS1, MV)	EXP(OLS)
RSP 0 a 0,1 m	ESF (OLS, WLS1, MV)	TODOS (OLS, WLS1, MV)	EXP (MV)
RSP 0,1 a 0,2 m	ESF (OLS,WLS1, MV) GAU (MV)	TODOS (OLS,WLS1, MV)	ESF (MV)
RSP 0,2 a 0,3 m	ESF (MV) GAU (OLS, WLS1)	NENHUM	GAU (MV)
UMD 0 a 0,1 m	GAU (OLS, MV) EXP(WLS1)	GAU (WLS1)	EXP (MV)
UMD 0,1 a 0,2 m	GAU (OLS, WLS1) EXP (MV)	ESF (WLS1)	EXP (OLS)
UMD 0,2 a 0,3 m	EXP (OLS) ESF (WLS1,MV)	NENHUM	ESF (MV)
PROD.	GAU (OLS, MV)	TODOS (OLS,WLS1, MV)	EXP (MV)

ESF: modelo esférico; EXP: modelo exponencial; GAU: modelo gaussiano; OLS: mínimos quadrados ordinários; WLS1: mínimos quadrados ponderados; MV: máxima verossimilhança.

dos casos, e o critério da validação cruzada foi considerado o mais seletivo. O critério de Akaike e do máximo valor do logaritmo da função verossimilhança não foram conclusivos. O critério da validação cruzada foi considerado o mais adequado para a escolha do melhor ajuste, e os mapas temáticos foram construídos utilizando-se uma estrutura de dependência espacial, escolhida segundo esse critério, que permitiu uma noção visual do comportamento dos atributos densidade, resistência do solo à penetração, umidade e produtividade na área estudada.

2. À semelhança da construção de modelos empíricos de variabilidade espacial aos semivariogramas experimentais, em que o conhecimento do pesquisador relativo às variáveis em estudo conduz a resultados melhores, também no caso de verificação de qualidade dos ajustes o conhecimento e o bom senso do pesquisador são de grande valia.

AGRADECIMENTOS

Ao CNPq, à Fundação Araucária e à Companhia Nacional de Abastecimento – CONAB, pelo apoio financeiro para a realização da pesquisa.

LITERATURA CITADA

AKAIKE, H. Information theory as an extension of the maximum likelihood principle. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2., Budapest, 1973. Proceedings. Budapest, Akadémia Kiadó, 1973. p.267-281.

COX, D.R. & HINKLEY, D.V. Theoretical statistics. London, Chapman and Hall, 1974. 532p.

CRESSIE, N. Fitting variogram models by weighted least squares. *Mathem. Geol.*, 17:563-586, 1985.

CRESSIE, N. Statistics for spatial data. New York, John Wiley, 1993. 900p.

CRESSIE, N. & HAWKINS, D.M. Robust estimation of the variogram. *Mathem. Geol.*, 12:115-125, 1980.

FILLIBEN, J.J. The probability plot correlation coefficient test for normality. *Technometrics*, 17:111-117, 1975.

GOMES, F.P. & GARCIA, C.H. Estatística aplicada a experimentos agrônômicos e florestais. Piracicaba, FEALQ, 2002. v.11. 305p.

ISAAKS, E.H. & SRIVASTAVA, R.M. An introduction to applied geostatistics. New York, Oxford University Press, 1989. 561p.

LEE, S.I. Validation of geoestatistical models using the Filliben test for orthonormal residual. *J. Hydrol.*, 158:319-332, 1994.

MARDIA, K.V. & MARSHALL, R.J. Maximum likelihood models for residual covariance in special regression. *Biometrika*, 71:135-146, 1984.

McBRATNEY, A. & WEBSTER, R. Choosing functions for semi-variograms os soil properties and fitting them to sample estimates. *J. Soil Sci.*, 37:617-639, 1986.

MELLO, J.M.; BATISTA, J.L.F.; RIBEIRO JR, P.J. & OLIVEIRA, M.S. Ajuste e seleção de modelos espaciais de semivariograma visando à estimativa volumétrica de *Eucalyptus grandis*. *Sci. For.*, 69:25-37, 2005.

- R DEVELOPMENT CORE TEAM (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<http://www.R-project.org>>. Acesso em 3 jun. 2005.
- RIBEIRO JR, P.J. & DIGGLE, P.J. geoR: A package for geostatistical analysis. R-NEWS, 01, 2001. Disponível em: <http://cran.r-project.org/doc/Rnews>.
- SAKAMOTO, Y.; ISHIGURO, M. & KITAGAWA, G. Akaike information criterion statistics. New York, Springer, 1986. 320p.