

Aplicação de modelos geostatísticos para o índice de cone

Ricardo Alves de Olinda - LCE, ESALQ/USP¹

Carlos Tadeu dos Santos Dias - LCE, ESALQ/USP²

Raphael Antonio Prado Dias - LCE, ESALQ/USP³

José Paulo Molin - LER, ESALQ/USP⁴

Paulo Justiniano Ribeiro Junior - LEG, UFPR/PR⁵

Abstract: This research aimed to study the spatial dependence of the penetration resistance of soil maps and evaluate conditional probabilities, finding a threshold resistance to root penetration and emergence of seeds. The geostatistical techniques are important tool for precision agriculture, because their use can cheapen the costs of monitoring soil compaction and simultaneously generate predictive maps of high quality.

Introdução

A resistência do solo à penetração do índice de cone (*IC*) tem sido frequentemente utilizado como indicador da compactação do solo em sistemas de manejo, por ser um atributo diretamente relacionado ao crescimento das plantas e de fácil e rápida determinação (TORMENA & ROLOFF, 1996). Ao analisar dados de resistência à penetração do *IC* mediante os métodos estatísticos clássicos são ignoradas as consequências da heterogeneidade espacial sobre a representatividade dos valores médios das amostras. No entanto, os métodos geostatísticos estão sendo utilizados para analisar tanto a dependência espacial como para interpolar atributos de solo através da krigagem (SOUZA et al., 2004). Dessa forma, o presente trabalho teve por objetivo estudar a dependência espacial da resistência a penetração do solo através de semivariograma e avaliar mapas de probabilidades condicionais, com intuito de encontrar um limiar a resistência de penetração das raízes e de emergência das sementes.

Material

O estudo foi realizado em uma área agrícola pertencente à Fundação ABC, em Castro, PR, nas coordenadas geográficas 24° 51' de latitude sul e 49° 55' de longitude oeste, com cerca de 23 ha e altitude média de 1030 metros acima do nível do mar. O solo foi classificado como um Latossolo Vermelho-Amarelo distrófico. A área vem sendo cultivada com milho e soja no verão e com aveia e trigo no inverno, desde 1985.

¹Doutorando em Estatística e Experimentação Agronômica: ricardoestatistico@usp.br

²Professor Titular da Universidade de São Paulo: ctdias@esalq.usp.br

³Doutorando em Estatística e Experimentação Agronômica: rapdias@esalq.usp.br

⁴Professor Associado da Universidade de São Paulo: jpmolin@esalq.usp.br

⁵Professor Adjunto da Universidade Federal do Paraná: paulojus@ufpr.br

Campos aleatórios

Um campo aleatório é um processo estocástico que existe em algum espaço real d -dimensional, definido por:

$$\{Z(s_i) : s_i \in T \subset \mathfrak{R}^d\}$$

em que, $Z(s_i)$ denota a variável aleatória Z na localização s_i do espaço sob uma região de estudo T . A descrição de um campo aleatório é obtida através das distribuições acumuladas finito-dimensionais F , para qualquer conjunto de localizações (s_1, s_2, \dots, s_n) pertencentes à região T e qualquer inteiro n :

$$F_{S_1, S_2, \dots, S_n}(z_1, z_2, \dots, z_n) \equiv P(Z(S_1) \leq z_1, Z(S_2) \leq z_2, \dots, Z(S_n) \leq z_n).$$

Semivariograma

A quantidade $\gamma(h)$ é conhecida como semivariância e é a metade da esperança da variância entre pares de pontos separados por uma distância h . Obedecendo-se às condições da hipótese de estacionariedade, a semivariância pode ser estimada simplesmente através de uma única realização do processo estocástico

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(s_i) - z(s_i + h)]^2,$$

em que $\gamma(h)$ é a semivariância estimada entre pares de pontos, $N(h)$ é o número de pares de valores medidos $z(s_i)$, $z(s_i + h)$, separados pela distância h .

Função de Matérn

Essa família de funções de correlação foi proposta por Bertil Matérn em 1986 e possui a seguinte forma:

$$\rho(h) = \{2^{k-1} \Gamma(\kappa)\}^{-1} (h/\phi)^\kappa K_\kappa(h/\phi),$$

sendo $\Gamma(\cdot)$ a função gama, $\kappa(\cdot)$ é a função Bessel de ordem κ , h a distância euclidiana entre duas localizações quaisquer do campo aleatório, os parâmetros dessa função são $\phi > 0$ e $\kappa > 0$.

Função Exponencial Potência

Essa família de funções de correlação possui características similares às da função de Matérn em que ϕ e κ são interpretados de forma análoga. Entretanto, neste caso, κ é limitado ao intervalo $[0; 2]$ e não definindo uma sequência de diferenciabilidade dos processos associados que passam de não diferenciáveis para $\kappa < 2$ para infinitamente diferenciáveis para $\kappa > 2$. A forma desta função de correlação é:

$$\rho(h) = \exp(h/\phi)^\kappa,$$

essas duas funções são muito utilizadas devido a capacidade de produzir comportamentos distintos quanto a suavidade do processo.

Modelos geoestatísticos gaussianos

Considerando que em alguma área T exista um campo aleatório gaussiano Z latente. Sendo assim, existe um vetor $Y(s)$ de $n \times 1$ valores observados em $s = (s_1, s_2, \dots, s_n)$, que segundo Diggle & Ribeiro Jr. (2007), sob o pressuposto de gaussianidade, pode ser modelado como:

$$Y(s) = \mu(x) + Z(s) + \varepsilon, \quad (1)$$

em que $\mu(x) = X\beta$ sendo X uma matriz $n \times q$ contendo $q - 1$ possíveis covariáveis, β um vetor $q \times 1$ de parâmetros associados a X , $Z(s)$ um campo aleatório gaussiano que possui vetor de médias $n \times 1$ nulo e matriz de covariâncias Σ de dimensão $n \times n$, sendo que cada elemento $\sum_{i,j}$ é igual a $Cov(Z(s_i); Z(s_j))$, para todo s_i e s_j pertencentes a s , e ε um vetor $n \times 1$ de ruídos brancos.

Estimação dos parâmetros

A expressão de uma estimativa empírica para o semivariograma é dada por:

$$\hat{\gamma} = \frac{\sum_{|N(h)|} (Z(s_i) - Z(s_j))^2}{2 |N(h)|} \quad (2)$$

em que, $|N(h)|$ é o número de pontos abrangidos pela distância h . Por simplicidade matemática, normalmente utiliza-se o logaritmo da função de verossimilhança para fazer a estimação, que associada a (1) tem a seguinte forma:

$$l(\theta; Y(s)) = -0.5(n \ln(2\pi) + \ln(|\Sigma_Y|) + (Y(s) - X\beta)^t \Sigma_Y^{-1} (Y(s) - X\beta)) \quad (3)$$

Krigagem

Suponha que o interesse é fazer predição para Z nas localizações s^* , a distribuição de probabilidade de $(Z(s^*)|Y(s))$ gera as predições com melhor precisão das estimativas. Dessa forma pode-se obter:

$$E(Z(s^*)|Y(s)) = \mu_Z + \Sigma_{Z,Y} \Sigma_Y^{-1} (Y(s) - X\beta) \quad (4)$$

sendo μ_Z a média de $Z(s^*)$ e $\Sigma_{Z,Y}$ é a matriz de covariâncias cruzadas entre $Z(s^*)$ e $Y(s)$. Além disso a variância preditiva é conhecida:

$$Var(Z(s^*)|Y(s)) = \Sigma_Z - \Sigma_{Z,Y} \Sigma_Y^{-1} \Sigma_{Z,Y}^t \quad (5)$$

sendo Σ_Z a matriz de covariâncias de $Z(s^*)$.

Resultados e Discursão

Na Figura 1, encontra-se o envelope simulado do variograma empírico utilizando 1000 simulações Monte Carlo, observa-se que vários pontos estão fora do envelope simulado indicando uma considerável dependência espacial no *IC*. Para que haja dependência, deve haver ao menos um ponto do variograma fora do envelope simulado, deste modo, se isso ocorrer rejeita-se a hipótese nula, de que não há dependência espacial (DIGGLE & RIBEIRO Jr., 2007).

As dependências espaciais observadas para os valores do *IC* mostram que as análises estatísticas clássicas, em que as observações são consideradas aleatórias e independentes espacialmente e que tratam o solo como um meio homogêneo, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas e que tratam o solo como um meio heterogêneo. De forma semelhante à observada por Tormena & Roloff 1996, com esse procedimento, é reduzido o risco de superdimensionar-se a operação de intervenção em alguns setores e subdimensionar-se em outros.

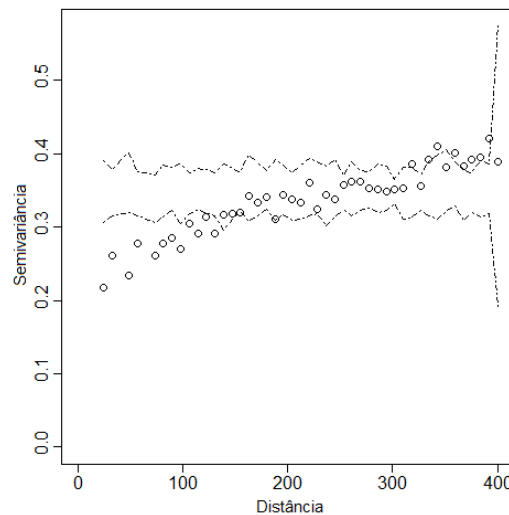


Figura 1: Verificando a dependência espacial através do envelope simulado.

As estimativas de máxima verossimilhança para os parâmetros do modelo univariado ajustado a compactação do solo (Tabela 1), considerando a matriz de delineamento X : média constante sob a região de estudo. As diferentes escolhas referentes ao κ produz diferença significativa na estimativa do parâmetro de variabilidade associado ao modelo. No entanto, o parâmetro ϕ é o mais afetado, o que é esperado, uma vez que, os parâmetros provenientes da função de correlação Matérn não são ortogonais. Analisando os valores maximizados do logaritmo da função de máxima verossimilhança e os seus respectivos *AIC* e *BIC*, foi selecionado o modelo exponencial potência com $\kappa=0,5$.

A análise variográfica apresentada na Figura 2 mostra que todas as características estudadas

Tabela 1: Estimativas de máxima verossimilhança dos parâmetros associados ao modelo para o índice de cone com média constante sobre a região de estudo.

$\hat{\beta}$	$\hat{\tau}^2$	$\hat{\sigma}^2$	$\hat{\phi}$	κ	$l(\hat{\theta})$	<i>AIC</i>	<i>BIC</i>
2,30	0,21	0,48	349,99	0,5	-285,5	578,97	594,50
3,10	0,25	2,75	349,95	1,5	-290,3	588,62	604,14
2,00	0,23	0,16	36,76	2,5	-286,0	579,91	595,43

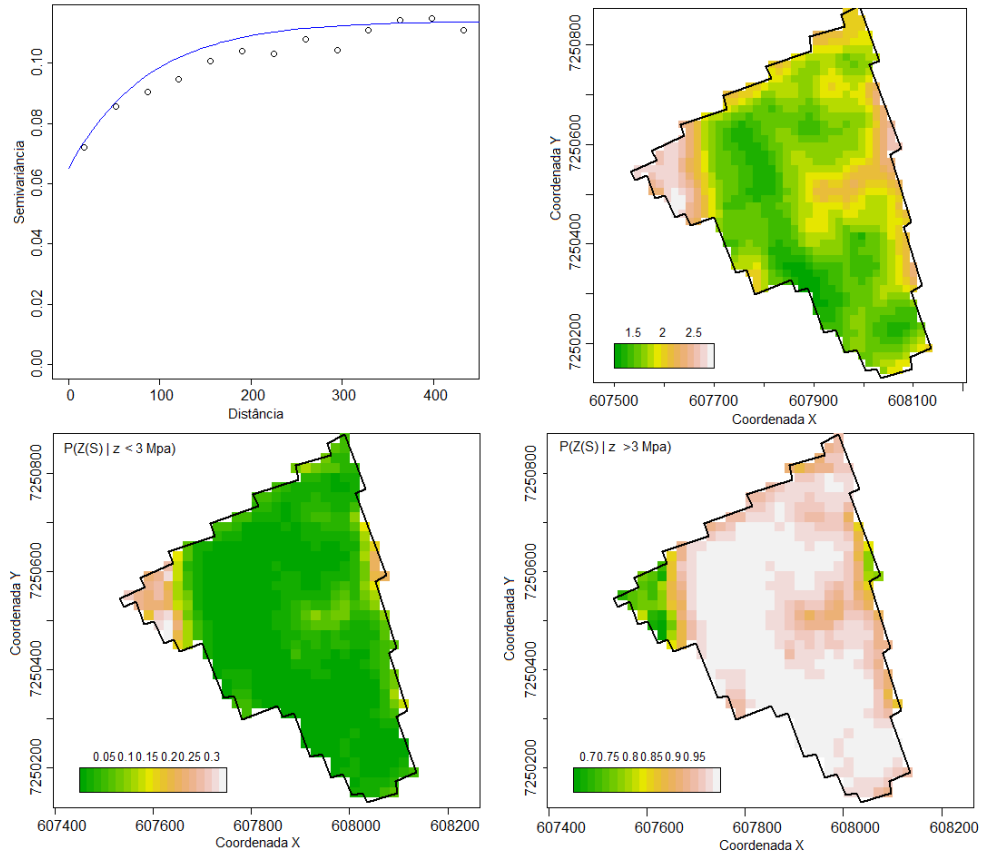


Figura 2: Gráficos de semivariograma, krigagem média e os mapas de probabilidade condicional.

nos procedimentos aqui descritos, apresentaram estruturas espaciais bem definidas, ou seja, existe uma função estrutural com semivariância de comportamento modelável. A aplicação do modelo gaussiano univariado permitiu que a variância espacial ficasse mais distante das amostras consideradas na interpolação. Esse modelo, associado a um efeito pepita (τ^2) baixo (0,06), permite que o campo estruturado tenha alta significância na interpolação.

Para verificar o grau de dependência espacial, entre as características dos índices de cone, utilizou-se da relação entre a variação estruturada e o patamar $\sigma^2 + \tau^2$. O $\kappa = 0,5$ foi o que apresentou o maior nível de dependência espacial. Pela classificação de Cambardella et al. (1994) pode ser classificada de moderada dependência espacial.

Com o modelo final estabelecido, foi calculado as predições espaciais para o *IC* em uma malha de 550 localizações espaciais, sendo que, as estimativas paramétricas foram substituídas nas

fórmulas de krigagem. A precisão da krigagem ficou num patamar aceitável e conseqüentemente tem-se boas informações sob a região de estudos (Figura 2).

Outra importante análise que se dispõe com a utilização dos mapas de krigagem são os mapas de probabilidades condicionais. Observa-se na Figura 2 uma grande área em que a compactação do solo é menor que 3MPa (Mega Pascal), ou seja, boas caracterizações do solo em termos de habilidade de crescimento das culturas. Apenas em pequenas regiões a resistência a penetração ultrapassa esse valor que é considerado como valor limite para resistência a penetração das raízes e de emergência das sementes (TAYLOR & GARDNER, 1963).

Conclusão

O uso das técnicas da geoestatística, tais como, semivariograma e a krigagem, são ferramentas imprescindíveis para a agricultura de precisão, pois sua utilização pode baratear os custos dos monitoramentos da compactação do solo e ao mesmo tempo gerar mapas preditivos de alta qualidade.

Referências Bibliográficas

- CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field scale variability of soil properties in central Iowa soils. **Soil Science Society of America Journal**, v.58, n.5, p.1501-1511, 1994.
- DIGGLE, P.J.; RIBEIRO Jr., P.J. **Model-Based geostatistics**. New York: Springer, 2007. 230p.
- SOUZA, Z.M. et al. Variabilidade espacial da estabilidade de agregados e matéria orgânica em solos de relevos diferentes. **Pesquisa Agropecuária Brasileira**, v.39, p.491-499, 2004.
- TORMENA, C.A.; ROLOFF, G. Dinâmica da resistência à penetração de um solo sob plantio direto. **Revista Brasileira de Ciência do Solo**, v.20, p.333-339, 1996.
- TAYLOR, H.M.; GARDNER, H.R. Penetration of cotton seedling tap roots as influenced by bulk density, moisture content and strength of soil. **Soil Science**, v.96, n.3, p.153-156, 1963.