

INFLUÊNCIA DE “OUTLIERS” NA ESTRUTURA DE VARIABILIDADE ESPACIAL DE DADOS DE UMIDADE VOLUMÉTRICA DO SOLO NA CAPACIDADE DE CAMPO

MIRITZ, Guilherme Kunde¹; AQUINO, Leandro Sanzi²; BLÖDORN, Rafael³; TIMM, Luís Carlos⁴; NEBEL, Álvaro Luiz Carvalho⁵

¹Graduando em Agronomia, FAEM (Faculdade de Agronomia Eliseu Maciel), UFPel (Universidade Federal de Pelotas), Campus Universitário – Caixa Postal 354 – CEP 96010-900-Capão do Leão.

Bolsista IC-CNPq, guilhermemiriz@hotmail.com; ²Doutorando em Agronomia – Solos, DS (Departamento de Solos)/FAEM/UFPel; ³Graduando em Engenharia Agrícola, Faculdade de Engenharia Agrícola, Universidade Federal de Pelotas. Bolsista PIBIC-CNPq; ⁴Professor Adjunto DER/FAEM/UFPel; ⁵Professor Adjunto IFSul-Pelotas

1 INTRODUÇÃO

O conhecimento dos atributos do solo é de grande importância para a tomada de decisão na busca de um manejo mais adequado dos sistemas de irrigação e drenagem (Bernardo et al., 2006). Contudo, os atributos do solo apresentam uma variabilidade no espaço, resultante de seus diferentes materiais de origem e diferentes processos de formação e intempéries e/ou efeitos de práticas de manejo adotadas (McGraw, 1994).

Para o estudo e compreensão dessa variabilidade, surgiram as ferramentas geoestatísticas que são um conjunto de técnicas que permitem identificar e mapear a estrutura de variabilidade espacial dos atributos do solo. Entretanto, na maioria dos estudos, a hipótese geoestatística de estabilidade dos momentos de primeira e segunda ordem é assumida, além da hipótese intrínseca de semelhança da distribuição dos dados a uma distribuição normal teórica (Vieira, 2000). Quando essa última não é atendida, algumas alternativas são sugeridas, como a transformação dos dados, o ajuste dos dados a equação de um plano, dentre outras (Landim, 2003); no entanto, esses procedimentos podem dificultar a interpretação dos resultados. Uma alternativa é a identificação da ocorrência de valores discrepantes na distribuição dos dados, conforme adotado por Gonçalves et al. (2005). Quando há ocorrência desse tipo de dado é prudente a repetição da amostragem e da análise laboratorial, ou até mesmo, uma análise criteriosa para verificar a ocorrência dessa informação. Porém, em muitos casos é feito o tratamento desses valores, pois o fator época de coleta pode influenciar no resultado da análise. Várias são as formas de tratar os valores discrepantes, como substituí-los pela média aritmética do conjunto dos dados, suprimi-los da série dos dados, ou também uma interpolação simples entre os valores vizinhos.

Assim, esse trabalho teve o objetivo de verificar o efeito do tratamento de valores discrepantes na estrutura de variabilidade espacial e na identificação de zonas homogêneas em um conjunto de dados referentes ao conteúdo de água no solo retido na tensão referente à capacidade de campo (θ_{cc}).

2 METODOLOGIA (MATERIAL E MÉTODOS)

Uma malha regular com um espaçamento de 10 x 10 m entre os pontos, totalizando 100 pontos, foi estabelecida em uma área da Estação Experimental de Terras Baixas da Embrapa Clima Temperado, Pelotas-RS. Em

cada ponto foram retiradas amostras de solo com estrutura preservada para a determinação da curva de retenção de água no solo (tensões de 0, 1, 6, 10, 33, 100 e 1.500 kPa), conforme metodologia descrita em Embrapa (1997).

Para o presente trabalho foi adotado o conteúdo de água no solo a base de volume em equilíbrio na tensão de 10 kPa como capacidade de campo (θ_{CC}). Esta série de dados foi submetida a uma análise exploratória para verificar se a distribuição seguia uma distribuição normal teórica por meio das medidas de posição (média e mediana), de dispersão (coeficiente de variação) e do formato da distribuição (assimetria e curtose). Para a identificação da ocorrência de valores discrepantes (“outliers”) foi construído o gráfico “Box Plot”. Os dados identificados como “outliers” foram substituídos por um valor igual à média aritmética de seus primeiros vizinhos posicionados na malha amostral, compondo, dessa forma, a série de dados θ_{CC} sem “outliers”, ou seja, uma nova variável.

Para a confecção dos semivariogramas e ajuste dos mesmos ao modelo teórico foi utilizado o software geoestatístico GS+, versão 7.0 (Gamma Design Software, 2004). A estrutura de variabilidade espacial de θ_{CC} foi avaliada por meio dos parâmetros do semivariograma teórico: efeito pepita (C_0), patamar (C_0+C), alcance (A) e o grau de dependência espacial (GDE), segundo a classificação proposta por Cambardella et al. (1994).

3 RESULTADOS E DISCUSSÕES

O gráfico “Box Plot” dos valores de θ_{CC} é apresentado na Figura 1. Analisando a figura é possível verificar a presença de valores discrepantes na distribuição dos dados desta variável, sendo um valor discrepante acima da cerca superior e três valores abaixo da cerca inferior. Esses valores podem estar influenciando o afastamento do conjunto dos dados de uma distribuição normal, o que poderá conduzir a erros na interpretação dos resultados quando da aplicação da ferramenta geoestatística, assumindo a hipótese intrínseca.

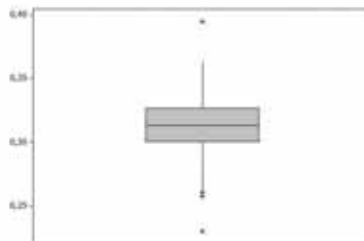


Figura 1 – Gráfico “Box Plot” ilustrando a presença de “outliers” na distribuição dos dados de θ_{CC} na malha experimental.

Na Tabela 1 são apresentados os resultados da estatística descritiva para os dados de θ_{CC} , na presença e ausência de “outliers”. Em ambos os casos verifica-se que os valores da média mantêm-se próximos aos da mediana, que o valor da assimetria permaneceu próximo a zero e que o valor da curtose tendeu a zero após a retirada dos “outliers”, indicando que a distribuição dos dados tendeu a uma mesocúrtica (coeficiente de curtose = 0). Em relação à dispersão dos dados em torno da média, expresso pelo coeficiente de variação (CV), verifica-se

que para ambas as situações, o CV é classificado como de baixa dispersão (CV<15%, Wilding & Drees, 1983).

Tabela 1. Estatística descritiva da variável θ_{CC} ($\text{cm}^3.\text{cm}^{-3}$) na presença e ausência de “outliers”.

Variáveis	Média ($\text{cm}^3.\text{cm}^{-3}$)	Mediana ($\text{cm}^3.\text{cm}^{-3}$)	CV (%)	Assimetria	Curtose
θ_{CC}	0,3118	0,3132	7,77	-0,2842	1,679
θ_{CC} sem outliers	0,3119	0,3132	6,62	-0,3021	-0,006

Na Figura 2 são apresentados os semivariogramas experimentais e teóricos da variável θ_{CC} na presença de “outliers” (Fig. 2-A) e na ausência (Fig. 2-B), e seus respectivos mapas de zonas homogêneas (Figs. 2-C e 2-D). Analisando os semivariogramas percebe-se que a faixa de dependência espacial dos valores de θ_{CC} diminuiu com a retirada de “outliers”, passando de 60,5 m para 44,4 m. Ambas as distribuições apresentam forte grau de dependência espacial (GDE<25%, Cambardella et al., 1994). Os parâmetros de ajuste do semivariograma não foram muito afetados, pois o coeficiente de determinação (r^2) passou de 0,956 para 0,962. Pelos mapas de contorno é possível verificar que a retirada de “outliers” do conjunto de dados propiciou um maior detalhamento na delimitação das zonas homogêneas de θ_{CC} .

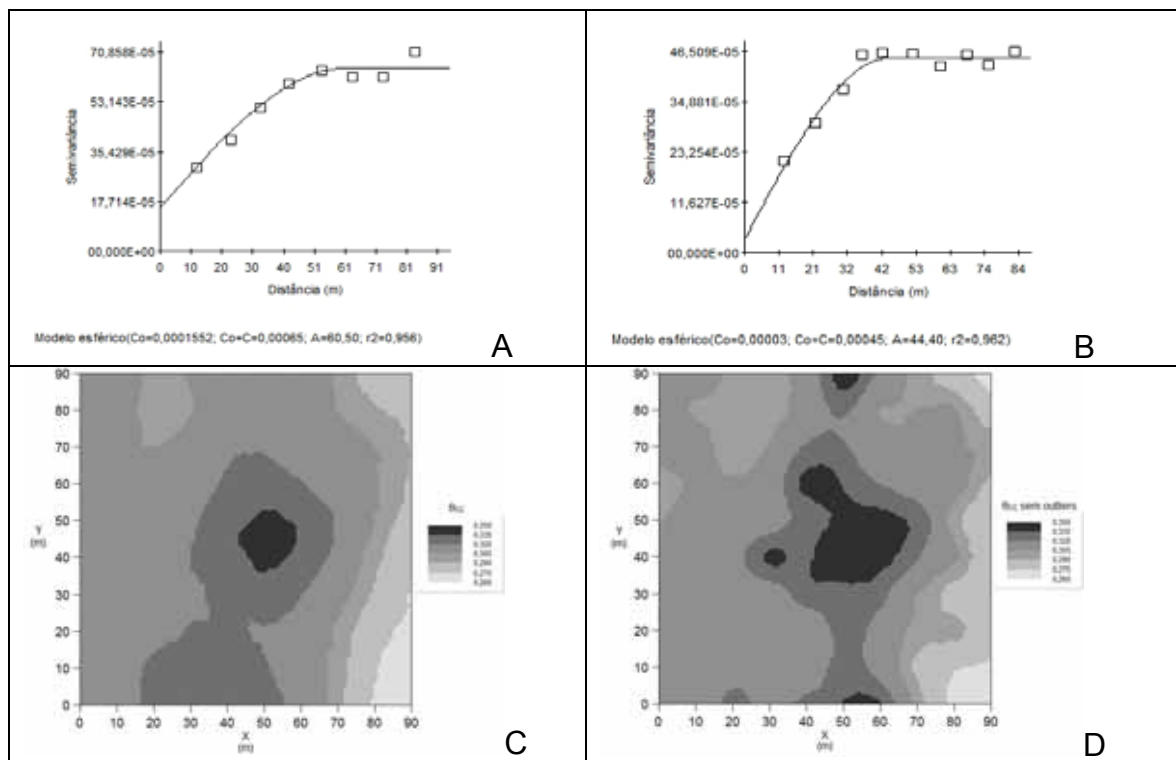


Figura 2 – Semivariogramas experimentais e teóricos da variável θ_{CC} , na presença de “outliers” (Fig. 2-A) e na ausência (Fig. 2-B) e seus respectivos mapas de zonas homogêneas (Figs. 2-C e 2-D).

4 CONCLUSÕES

O tratamento dos valores discrepantes na distribuição dos dados de conteúdo de água no solo referente à capacidade de campo (θ_{CC}) alterou a sua estrutura de variabilidade reduzindo a sua faixa de dependência espacial de 60,50 m para 44,40 m. Além disso, propiciou um maior detalhamento na delimitação das zonas homogêneas de θ_{CC} na área experimental.

5 AGRADECIMENTOS

A Embrapa Clima Temperado pela concessão da área experimental. Ao CNPq pela concessão dos recursos financeiros para o desenvolvimento do projeto e pela concessão de bolsas de doutorado e de iniciação científica.

6 REFERÊNCIAS

- BERNARDO, S.; SOARES, A.A.; MANTOVANI, E.C. **Manual de irrigação**. 8^a ed. atualizada e ampliada. Viçosa: Editora UFV, 2006. 625p.
- 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, p. 1501-1511, 1994.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. Centro Nacional de Pesquisa de Solos. **Manual de Métodos de Análise de Solo**. 2^a ed. Rio de Janeiro: Embrapa Solos, 1997. 212 p.
- GAMMA DESIGN SOFTWARE. GS+: **Geostatistics for the Environmental Sciences**. Plainwell: Gamma Design Software, 2004.
- GONÇALVES, M.G.; FARIAS, L.N.; COSTA, A.D.; MOTA, L.S.; OLIVEIRA, B.L.; CEDIA, M.B.; VIEIRA, S.R. Contribuição da geoestatística para o planejamento experimental de uma pastagem do sistema integrado de produção agroecológica (Fazendinha km-47). **Revista Universidade Rural, Série Ciências da Vida**, v. 25, p. 1 – 5, 2005.
- LANDIM, P.M.B. **Análise estatística de dados geológicos**. 2^o edição. São Paulo: UNESP, 2003. 253p.
- McGRAW, T. Soil test level variability in Southern Minnesota. **Better Crops with Plant Foods**, v. 78, p. 24-25, 1994.
- VIEIRA, S.R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R.F.; ALVAREZ, V.H. & SCHAEFER, C.E.G.R. (Eds.). **Tópicos em ciência do solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. v.1. p.1-54..
- WILDING, L.P. & DREES, L.R. Spatial variability and pedology. In: WILDING, L. P.; SMECK, N.E. & HALL, G.F. (Eds.) **Pedogenesis and soil taxonomy: concepts and interactions**. New York: Elsevier, 1983. p.83-116.