

Proposta de uma generalização para os modelos de semivariogramas Exponencial e Gaussiano

Proposal of a generalization for Exponential and Gaussian semivariogram models

Enio Júnior Seidel¹; Marcelo Silva de Oliveira²

Resumo

O objetivo do trabalho foi propor uma correção para os modelos exponencial e gaussiano em função de diferentes percentuais de explicação da contribuição. O procedimento consistiu em determinar o valor de uma correção, denominada de k , com base no percentual de explicação da contribuição que se deseja atingir. Foi calculada a correção k para 95% e 99,99% da contribuição, mostrando ser possível obter diferentes expressões matemáticas para os modelos exponencial e gaussiano. Além disso, foram construídas as expressões generalizadas para esses dois modelos. Para uma melhor observação do comportamento dos modelos, para diferentes valores de k , é simulado um cenário de dependência espacial e, a partir desse cenário, são ajustados os modelos exponencial e gaussiano considerando o alcance obtido a 95% e 99,99% da contribuição. Foi possível realizar a correção, e, com base nessa, construiu-se uma generalização dos modelos exponencial e gaussiano. Além disso, visualizou-se que é possível modelar diferentes dependências espaciais, pois se podem considerar diferentes percentuais de explicação da contribuição. Contudo, é necessário lembrar que o percentual de 99,99% de explicação da contribuição é o mais próximo da realidade, evidenciando que a correção, com $k = 9$, é a mais ideal para uma maior aproximação do real comportamento do fenômeno.

Palavras-chave: Análise Variográfica. Continuidade espacial. Alcance de dependência espacial.

Abstract

The aim of this work was to propose a correction for Exponential and Gaussian models due to the different percentages of explanation of the contribution. The procedure consisted in determining the value of a correction called k , based on the percentage of explanation of the contribution that one wants to reach. Correction k was calculated for 95% and 99.99% of the contribution, showing that it is possible to obtain different mathematical expressions for Exponential and Gaussian models. In addition, generalized expressions for these two models were constructed. For better observation of the behavior of models for different values of k , a scenario of spatial dependence was simulated, and from this scenario, the Exponential and Gaussian models were fitted considering the range obtained at 95% and 99.99% of the contribution. It was possible to perform the correction, and based on the results, a generalization of Exponential and Gaussian models was constructed. Furthermore, it was possible to visualize that it is possible to model different spatial dependences, since different percentages of explanation of the contribution can be considered. However, it is noteworthy that the percentage of 99.99% of explanation of the contribution is the closest to reality, showing that the correction with $k = 9$ is the ideal situation for a better approximation to the actual behavior of the phenomenon.

Keywords: Variographic analysis. Spatial continuity. Range of spatial dependence.

¹ Docente da Universidade Federal do Pampa - UNIPAMPA, Campus Itaqui-RS; ejrseidel@hotmail.com.

² Docente do Departamento de Ciências Exatas da Universidade Federal de Lavras - UFLA; marcelo.oliveira@dex.ufla.br.

Introdução

O semivariograma é a principal ferramenta utilizada para estudar a estrutura de dependência espacial em Geoestatística. É um gráfico que relaciona semivariâncias (γ) com distâncias (h). Mais detalhes sobre o semivariograma e sua construção ver Journel e Huijbregts (2003) e Olea (2006).

Após a construção do semivariograma é necessário ajustar um modelo teórico que, de acordo com Carvalho, Silveira e Vieira (2002), permite visualizar a natureza da variação espacial das variáveis estudadas, explicando o comportamento da dependência espacial. Os modelos mais utilizados na literatura são os modelos Esférico, Exponencial e Gaussiano. Esses modelos, na sua forma tradicional, possuem 4 parâmetros, que são o alcance, o patamar, o efeito pepita e a contribuição.

O alcance (a) é compreendido como a distância dentro da qual as amostras apresentam-se correlacionadas espacialmente. Amostras separadas por distâncias menores que o alcance são correlacionadas umas às outras, o que permite que se façam interpolações para espaçamentos menores do que os amostrados (VIEIRA, 2000).

O efeito pepita (C_0) é o valor que representa a descontinuidade do semivariograma, no sentido de que, à medida que h tende a zero, $\gamma(h)$ tende a um valor positivo C_0 . Wu et al. (2008) definem que o efeito pepita representa o erro experimental e a variação no espaço dentro de distâncias mínimas de amostragem. O efeito pepita é uma espécie de variância casual, isto é, é um tipo de variabilidade devida ao acaso.

A contribuição (C_1) é a diferença entre o patamar e o efeito pepita, sendo o segmento gráfico nesse intervalo que, praticamente, caracteriza a dependência espacial de processos estocásticos contínuos. A contribuição pode ser entendida como uma variância estruturada que é devida à estrutura espacial do fenômeno em estudo.

O patamar ($C = C_0 + C_1$) é a soma da contribuição, que é a variância estrutural, com o efeito pepita, que é a variância ao acaso. Assim, tem-se que o patamar pode ser entendido como a variância total, que, como observado em Vieira (2000) e Abreu et al. (2003), é aproximadamente igual à variância dos dados.

Contudo, necessita-se fazer uma observação em relação ao parâmetro alcance. Nos fenômenos em que se ajusta o modelo esférico, o alcance real é finito (a). Já nos fenômenos em que se ajustam o modelo exponencial ou gaussiano, o alcance real é infinito. Contudo, para fins de operações matemáticas, convencionou-se que o alcance, nos modelos exponencial e gaussiano, é finito (respectivamente, $a/3$ e $a/\sqrt{3}$). Esses valores virtuais de alcance são devidos ao fato de que, nos modelos exponencial e gaussiano, o alcance (alcance prático) é obtido a 95% da contribuição como observado por Biondi, Myers e Avery (1994).

Porém, na literatura clássica (ISAAKS; SRIVASTAVA, 1989; JOURNEL; HUIJBREGTS, 2003; WACKERNAGEL, 2003; WEBSTER; OLIVER, 2007) não se tem uma descrição ou demonstração do porque utilizar-se o fator 95% da contribuição. Dessa forma, tem-se o objetivo de propor uma correção para os modelos exponencial e gaussiano em função de diferentes percentuais de explicação da contribuição. Tal correção possibilita a construção de uma nova classe (generalização) de modelos exponenciais e gaussianos.

Metodologia

As expressões matemáticas dos modelos esférico, exponencial e gaussiano, da forma utilizada atualmente, são apresentadas a seguir, de acordo com Issaks e Srivastava (1989). O modelo esférico é dado por:

$$\gamma(h)_{esf} = \begin{cases} 0 & , h = 0 \\ C_0 + C_1 \left[1,5 \left(\frac{h}{a} \right) - 0,5 \left(\frac{h}{a} \right)^3 \right] & , 0 < h \leq a \\ C_0 + C_1 & , h > a \end{cases} \quad (1)$$

O modelo exponencial é dado por:

$$\gamma(h)_{exp} = \begin{cases} 0 & , h = 0 \\ C_0 + C_1 \left[1 - e^{-\left(\frac{h}{a/3} \right)} \right] & , h \neq 0 \end{cases} \quad (2)$$

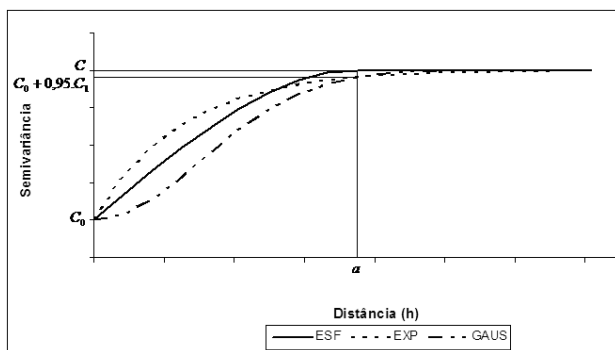
O modelo gaussiano é dado por:

$$\gamma(h)_{gau} = \begin{cases} 0 & , h = 0 \\ C_0 + C_1 \left[1 - e^{-\left[\left(\frac{h}{a/\sqrt{3}} \right)^2 \right]} \right] & , h \neq 0 \end{cases} \quad (3)$$

em que C_0 é o efeito pepita, C_1 é a contribuição, a é o alcance (alcance prático) e h é a distância entre observações.

A Figura 1 apresenta o comportamento gráfico dos modelos esférico (ESF), exponencial (EXP) e gaussiano (GAUS). Em destaque, a Figura 1 mostra, graficamente, que, nos modelos exponencial e gaussiano, o alcance é atingido em 95% da contribuição.

Figura 1 - Comportamento gráfico dos modelos esférico (ESF), exponencial (EXP) e gaussiano (GAUS).



Fonte: os autores.

O procedimento consiste em determinar o valor de uma correção, denominada de k , para os modelos exponencial e gaussiano, com base no percentual de explicação da contribuição que se deseja atingir.

É calculada a correção k para 95% e 99,99% da contribuição, mostrando que é possível obter diferentes expressões matemáticas para os modelos exponencial e gaussiano.

Para uma melhor observação do comportamento dos modelos para diferentes valores de k , é simulado um cenário, de dependência espacial, que possui $C_0=12,5$, $C_1=37,5$ e $a=50$. Além disso, esse cenário, simulado com o pacote *RandomFields* (SCHLATHER, 2001), do software R (R CORE TEAM, 2012), apresenta as seguintes características:

- 1º) média=0;
- 2º) patamar=50;
- 3º) $n=169$;
- 4º) *grid* regular de 100x100m;
- 5º) processo estocástico gaussiano
- 6º) estimador clássico de semivariograma;

A partir desse cenário são ajustados os modelos exponencial e gaussiano considerando o alcance obtido a 95% e 99,99% da contribuição. Os ajustes são realizados utilizando-se o pacote *geoR* (RIBEIRO JUNIOR; DIGGLE, 2001) do software R (R CORE TEAM, 2012).

Resultados e discussão

A correção que se propõe, nos modelos exponencial e gaussiano, é tal que o alcance prático pode ser definido como aquele que é atingido a um percentual qualquer da contribuição (C_1), que é definido pelo pesquisador. Por exemplo, ao considerar 99,99% da contribuição (0,9999), consegue-se atingir maior proximidade do alcance real ($a = \infty$), se comparada com a modelagem realizada atualmente com 95% da contribuição,

conforme mostram as equações (2) e (3). Assim, observa-se que o alcance, no caso de se considerar 99,99% da contribuição, tem maior aproximação do alcance real (que é infinito), possuindo características mais adequadas para representar melhor o comportamento verdadeiro da dependência espacial do fenômeno em estudo. Deve-se ressaltar que k não se constitui em mais um parâmetro estatístico a ser estimado, mas sim, que é um novo coeficiente que também pode controlar a modelagem dos semivariogramas exponencial e gaussiano sobre nuvens de pontos originadas de dados de observação.

Ao considerar 99,99% da contribuição como sendo o valor para o qual se obtém o valor do alcance prático para o modelo exponencial, tem-se que a correção (k) é igual a 9. Assim, de posse deste novo valor, é possível reescrever a equação do modelo exponencial como:

$$\gamma_{\text{exp}}(h) = C_0 + C_1 \left[1 - e^{\left(-\frac{h}{a/9} \right)} \right] \quad (4)$$

em que C_0 é o efeito pepita, C_1 é a contribuição, a é o alcance (alcance prático) e h é a distância entre observações.

As Figuras 2 e 3 mostram, respectivamente, o desenvolvimento do cálculo da correção k , para 95% e para 99,99% da contribuição, nos modelos exponencial e gaussiano.

Figura 2 - Cálculo da correção para os modelos exponencial e gaussiano considerando 95% da contribuição.

Exponencial:	Gaussiano:
$\Rightarrow \gamma_{\text{exp}}(h) = C_0 + 0,95.(C_1)$	$\Rightarrow \gamma_{\text{gaus}}(h) = C_0 + 0,95.(C_1)$
$\Rightarrow C_0 + C_1 \left[1 - e^{\left(-\frac{a}{a/\sqrt{k}} \right)} \right] = C_0 + 0,95.(C_1)$	$\Rightarrow C_0 + C_1 \left[1 - e^{\left[-\left(\frac{a}{a/\sqrt{k}} \right)^2 \right]} \right] = C_0 + 0,95.(C_1)$
$\Rightarrow 1 - e^{\left(-\frac{a}{a/\sqrt{k}} \right)} = 0,95$	$\Rightarrow 1 - e^{\left[-\left(\frac{a}{a/\sqrt{k}} \right)^2 \right]} = 0,95$
$\Rightarrow e^{(-k)} = 0,05$	$\Rightarrow e^{(-k)} = 0,05$
$\Rightarrow -k = \ln(0,05)$	$\Rightarrow -k = \ln(0,05)$
$\Rightarrow k = 2,996$	$\Rightarrow k = 2,996$
$\Rightarrow k \cong 3$	$\Rightarrow k \cong 3$

Fonte: os autores.

Figura 3 - Cálculo da correção para os modelos exponencial e gaussiano considerando 99,99% da contribuição.

Exponencial:	Gaussiano:
$\Rightarrow \gamma_{\text{exp}}(h) = C_0 + 0,9999.(C_1)$	$\Rightarrow \gamma_{\text{gaus}}(h) = C_0 + 0,9999.(C_1)$
$\Rightarrow C_0 + C_1 \left[1 - e^{\left(-\frac{a}{a/\sqrt{k}} \right)} \right] = C_0 + 0,9999.(C_1)$	$\Rightarrow C_0 + C_1 \left[1 - e^{\left[-\left(\frac{a}{a/\sqrt{k}} \right)^2 \right]} \right] = C_0 + 0,9999.(C_1)$
$\Rightarrow 1 - e^{\left(-\frac{a}{a/\sqrt{k}} \right)} = 0,9999$	$\Rightarrow 1 - e^{\left[-\left(\frac{a}{a/\sqrt{k}} \right)^2 \right]} = 0,9999$
$\Rightarrow e^{(-k)} = 0,0001$	$\Rightarrow e^{(-k)} = 0,0001$
$\Rightarrow -k = \ln(0,0001)$	$\Rightarrow -k = \ln(0,0001)$
$\Rightarrow k = 9,2$	$\Rightarrow k = 9,2$
$\Rightarrow k \cong 9$	$\Rightarrow k \cong 9$

Fonte: os autores.

No modelo gaussiano, ao se adotar 99,99% da contribuição, tem-se valor da correção k igual a 9, da mesma forma que no modelo exponencial. Portanto, é possível reescrever a equação do modelo gaussiano como:

$$\gamma_{\text{gaus}}(h) = C_0 + C_1 \left\{ 1 - e^{\left[-\left(\frac{h}{a/\sqrt{9}} \right)^2 \right]} \right\} \quad (5)$$

em que C_0 é o efeito pepita, C_1 é a contribuição, a é o alcance (alcance prático) e h é a distância entre observações.

Como tomou-se, por exemplo, 99,99% da contribuição para se obter o alcance (a), teve-se como valor da correção $k = 9$. Contudo, ao definir a correção como um valor geral k , tem-se uma generalização dos modelos exponencial e gaussiano. Essa generalização permite escrever as equações, respectivamente, nas formas:

$$\gamma_{\text{exp}}(h) = C_0 + C_1 \left[1 - e^{-\left(\frac{h}{a/k}\right)} \right] \quad (6)$$

e

$$\gamma_{\text{gaus}}(h) = C_0 + C_1 \left\{ 1 - e^{-\left[\left(\frac{h}{a/\sqrt{k}}\right)^2 \right]} \right\} \quad (7)$$

em que C_0 é o efeito pepita, C_1 é a contribuição, a é o alcance (alcance prático), h é a distância entre observações e k é a correção relacionada com o percentual da contribuição para o qual se define o alcance prático.

Na Tabela 1 apresenta-se uma relação entre o percentual da contribuição e o valor da correção k . Essa relação é proporcional, no sentido de que, quanto maior a explicação que se deseja obter da contribuição, maior é o valor da correção.

Tabela 1 - Relação entre a correção k , dos modelos exponencial e gaussiano e o percentual da contribuição que se deseja atingir.

Percentual da contribuição	Correção k
100%	∞
99,99%	9
⋮	⋮
95%	3
⋮	⋮
90%	2
⋮	⋮
0%	0

Fonte: os autores.

Os modelos exponencial e gaussiano, em suas formas generalizadas, respeitam as seguintes suposições para um semivariograma isotrópico e estacionário de segunda ordem, na forma em que são dadas em Schabenberger e Gotway (2005):

1ª) Se $\gamma(h)$ é válido no R^d , é também válido no R^s , com $s < d$;

2ª) Se $\gamma_1(h)$ e $\gamma_2(h)$ são semivariogramas válidos, então $a.\gamma_1(h) + b.\gamma_2(h)$, com $a, b \geq 0$, é um semivariograma válido;

3ª) Um semivariograma válido $\gamma(h)$ é condicionalmente negativa definida, isto é, $2 \sum_{i=1}^m \sum_{j=1}^m a_i a_j \gamma(s_i - s_j) \leq 0$, para quaisquer números

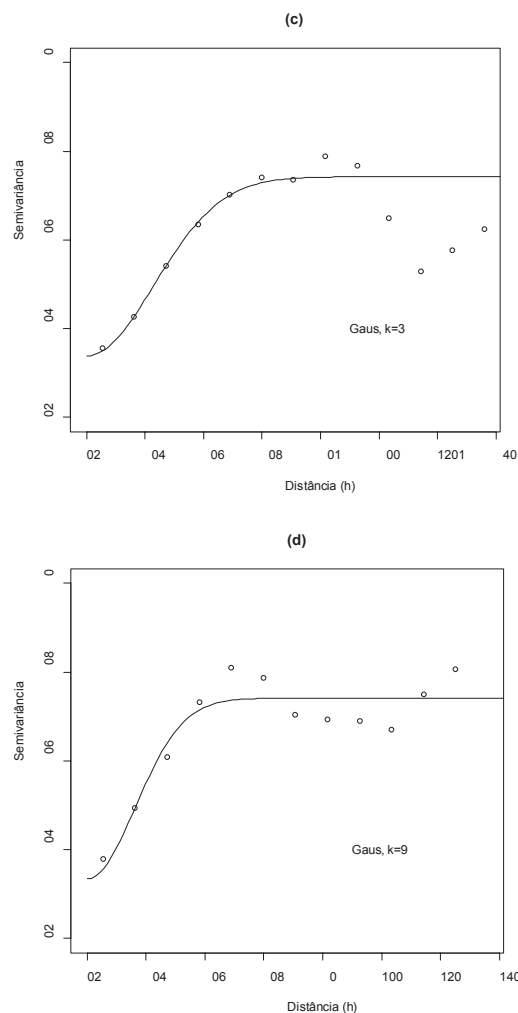
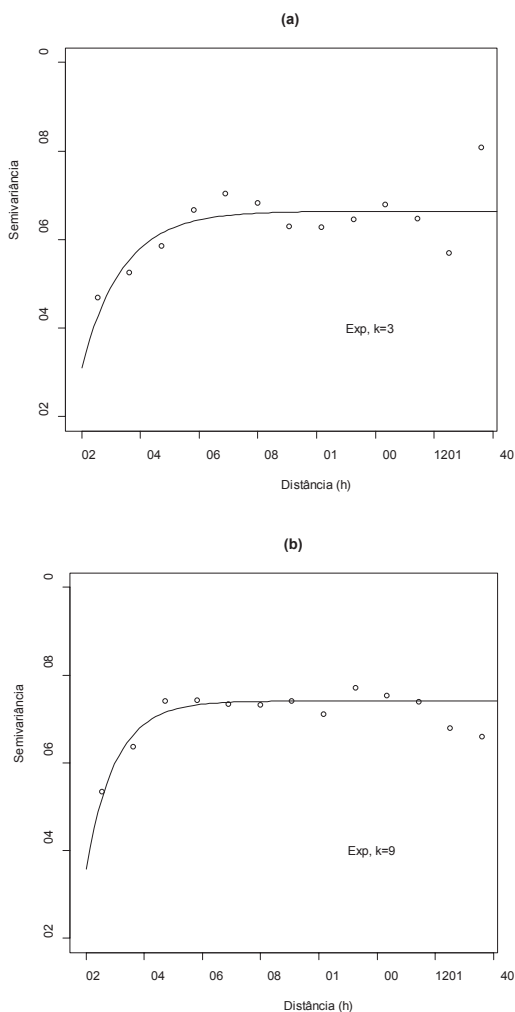
reais a_1, \dots, a_m tais que $\sum_{i=1}^m a_i = 0$ e qualquer número finito de locais;

4ª) Uma condição necessária para $\gamma(h)$ ser um semivariograma válido é a de que $2.\gamma(h)$ cresce mais lentamente que $\|h\|^2$.

Vê-se que a determinação de uma correção, a partir de um percentual da contribuição, revela a verdadeira natureza dos modelos exponencial e gaussiano. De fato, a consideração desse percentual permite criar uma classe de modelos novos para semivariogramas, permitindo ampliar o leque de modelagem de fenômenos exponenciais ou gaussianos.

O pesquisador tem, então, a opção de definir o valor de k a partir do percentual da contribuição que deseja atingir, ou explicar. A Figura 4 mostra os semivariogramas, a partir do cenário simulado, para os modelos exponencial e gaussiano. Esse cenário, no modelo exponencial, foi simulado com base na relação $a/3$ para 95% da contribuição e com base na relação $a/9$ para 99,99% da contribuição. Já no modelo gaussiano, o cenário foi simulado com base nas relações $a/\sqrt{3}$ e $a/\sqrt{9}$.

Figura 4 - Semivariogramas ajustados, por modelo exponencial com (a), por modelo exponencial com (b), por modelo gaussiano com (c) e por modelo gaussiano com (d), a partir do cenário simulado.



Fonte: os autores.

Observa-se, na Figura 4, o efeito da mudança de k : há um crescimento mais acentuado da nuvem de pontos quando k aumenta, mostrando que o ajuste de k pode fazer os modelos exponencial e gaussiano ajustarem-se a diferentes nuvens, que representam diferentes estruturas de dependência espacial. Esta é a generalização implicada pela correção utilizando o coeficiente k .

Conclusões

Foi possível realizar uma correção, denominada de k , nos semivariogramas exponencial e gaussiano a partir do conceito sobre o percentual de explicação da contribuição. A partir disso, construiu-se uma generalização dos modelos exponencial e gaussiano.

É possível modelar diferentes dependências espaciais, pois se podem considerar diferentes percentuais de explicação da contribuição. Contudo, é necessário lembrar que o percentual de 99,99% de explicação da contribuição é o mais próximo da realidade (mais próximo de 100%), evidenciando que a correção, com $k = 9$, é a mais ideal para uma maior aproximação do real comportamento do fenômeno.

Sugere-se para estudos futuros a ampliação do número de simulações, que foi uma das restrições do trabalho, e a implementação desses modelos generalizados em um *software*.

Agradecimento

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro em forma de bolsa.

Referências

ABREU, S. L.; REICHERT, J. M.; SILVA, V. R.; REINERT, D. J.; BLUME, E. Variabilidade espacial de propriedades físico-hídricas do solo, da produtividade e da qualidade de grãos de trigo em Argissolo Franco Arenoso sob plantio direto. *Ciência Rural*, Santa Maria, v. 33, n. 2, p. 275-282, 2003.

BIONDI, F.; MYERS, D. E.; AVERY, C. C. Geostatistically modeling stem size and increment in an old-growth forest. *Canadian Journal of Forest Research*, Ottawa, v. 24, n. 7, p. 1354-1368, 1994.

CARVALHO, J. R. P.; SILVEIRA, P. M.; VIEIRA, S. R. Geoestatística na determinação da variabilidade espacial de características químicas do solo sob diferentes preparos. *Pesquisa Agropecuária Brasileira*, Rio de Janeiro, v. 37, n. 8, p. 1151-1159, 2002.

ISAACS, E. H.; SRIVASTAVA, R. M. *Applied geostatistics*. New York: Oxford University Press, 1989.

JOURNAL, A. G.; HUIJBREGTS, C. J. *Mining geostatistics*. Caldwell: Blackburn Press, 2003.

OLEA, R. A. A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment*, Berlin, v. 20, n. 5, p. 307-318, 2006.

R CORE TEAM. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2012.

RIBEIRO JUNIOR., P. J.; DIGGLE, P. J. geoR: a package for geostatistical analysis. *R-NEWS*, Vienna, v. 1, n. 2, p. 15-18, 2001.

SCHABENBERGER, O.; GOTWAY, C. A. *Statistical methods for spatial data analysis*. Boca Raton: Chapman & Hall/CRC, 2005.

SCHLATHER, M. Simulation and analysis of random fields. *R- NEWS*, Vienna, v. 1, n. 2, p. 18-20, 2001.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. *Tópicos em ciência do solo*, Viçosa, v. 1, p. 1-53, 2000.

WACKERNAGEL, H. *Multivariate geostatistics: an introduction with applications*. New York: Springer-Verlag, 2003.

WEBSTER, R.; OLIVER, M. *Geostatistics for environmental scientists*. 2. ed. Chichester: Wiley, 2007.

WU, C.; WU, J.; LUO, Y.; ZHANG, H.; TENG, Y. Statistical and geoestatistical characterization of heavy metal concentrations in a contaminated area taking into account soil map units. *Geoderma*, Amsterdam, v. 144, p. 171-179, 2008.

Recebido em 16 Janeiro 2013 - Received on January 16, 2013.
Aceito em 13 Março, 2013 - Accepted on March 13, 2013.

