

## CARACTERIZAÇÃO DA VARIABILIDADE ESPACIAL DE PARÂMETROS HIDRÁULICOS EM SOLOS ALUVIAIS NO ESTADO DE PERNAMBUCO

**Suzana M. G. L. Montenegro**

UFPE. Departamento de Engenharia Civil. Recife, PE  
44smglm@npd.ufpe.br

**Abelardo A. A. Montenegro**

UFRPE. Departamento de Tecnologia Rural. Recife, PE  
monte@hotlink.com.br

**Rae Mackay**

School of Earth Sciences. University of Birmingham. Inglaterra  
r.mackay@bham.ac.uk

---

### RESUMO

A variabilidade da condutividade hidráulica saturada é analisada através das estatísticas clássica e geoestatística, destacando-se a importância dessa última para inferir a estrutura de correlação do parâmetro. Dados de uma área irrigada na região agreste do Estado de Pernambuco são utilizados. A estrutura de correlação identificada, compõe o modelo de variabilidade espacial do parâmetro, necessário à modelagem espacial estocástica de escoamento em solos não saturados, e é utilizada para interpolação, através de kriging, do parâmetro hidráulico em pontos não cobertos pela amostragem. Além da validação do modelo de variabilidade espacial por procedimentos geoestatísticos, destaca-se a importância da análise funcional, baseada em resultados da modelagem estocástica de fluxo.

### INTRODUÇÃO

Diversas são as linhas de pesquisa e estudos em recursos hídricos em que é necessário a caracterização da variabilidade espacial de parâmetros e variáveis envolvidas. Por exemplo, o conhecimento da variabilidade espacial é necessário ao simples cálculo da precipitação média sobre uma área.

Os solos são formações naturais reconhecidamente heterogêneas (Nielsen et al., 1973). Russo e Bresler (1981) demonstraram que os parâmetros relacionados à água no solo são altamente variáveis, inclusive em pequenas áreas. Cambardella et al. (1994), apontam que as distribuições espaciais das propriedades dos solos nas escalas de campo e da bacia, podem afetar respostas hidrológicas e de transporte de solutos (e. g. sais, herbicidas) na subsuperfície e na água subter-

rânea. O reconhecimento desse fato, segundo os autores, é de fundamental importância para o refinamento de práticas de manejo da agricultura e para avaliar os efeitos da agricultura na qualidade do meio ambiente.

Para consideração da heterogeneidade, ou variabilidade espacial das propriedades do solo, na modelagem de escoamento e transporte de poluentes em meios porosos, duas abordagens diferentes podem ser consideradas: a determinística e a estocástica. A abordagem determinística assume que os parâmetros possuem valores definidos em cada ponto do espaço, e quanto a abordagem estocástica reconhece a incerteza nos valores dos parâmetros e, por consequência das variáveis, associa a ambos erros de determinação e variabilidade espacial.

Para consideração da variabilidade espacial na abordagem determinística é necessário geralmente um grande número de pontos de observação do parâmetro em questão. Além disso, normalmente a modelagem determinística requer o conhecimento do parâmetro em pontos não amostrados. A análise estocástica, por sua vez, requer a representação da incerteza dos parâmetros para a geração de possíveis cenários de distribuição espacial desses. A necessidade de interpolar e extrapolar valores de quantidades físicas, através de kriging por exemplo, para fins de modelagem, induz a necessidade de se estabelecer padrões de variabilidade espacial.

A geoestatística, ou Teoria das Variáveis Regionalizadas (Matheron, 1963), é a ciência que estuda a variabilidade de uma quantidade através de um modelo probabilístico, associado à posição relativa das observações no espaço e possíveis correlações dessas observações em função da distância de separação entre elas. O modelo probabilístico isoladamente, base da

estatística clássica que considera a quantidade como variável aleatória, não é capaz de identificar essa correlação espacial, através da qual as observações estão tanto mais correlacionadas entre si quanto menor for a distância de separação. Essa estrutura de correlação é a base para métodos de interpolação espacial, como por exemplo o kriging. Na metodologia geoestatística a quantidade não mais é identificada como uma variável aleatória, uma vez que essa estrutura de correlação é levada em consideração, mas sim como uma Variável Regionalizada (VR).

O objetivo deste estudo é o uso da geoestatística para a caracterização da variabilidade espacial da condutividade hidráulica saturada ( $K_{sat}$ ), ou seja, a condutividade hidráulica de solos não saturados na condição de umidade de saturação, para fins de modelagem estocástica de escoamento.

## METODOLOGIA GEOESTATÍSTICA

### Generalidades

A metodologia geoestatística é baseada na hipótese de ergodicidade, através da qual o modelo de probabilidade do conjunto de possíveis realizações da variável pode ser inferido a partir de uma única realização, ou seja, a partir do conjunto de dados experimentais. Essa hipótese estabelece uma equivalência entre variabilidade espacial e incerteza, por permitir um intercâmbio entre a média (valor esperado) espacial e a média do conjunto de possíveis realizações. Para invocar essa hipótese, é necessário assumir que o campo é estatisticamente homogêneo, ou estacionário, implicando que a função de probabilidade descrevendo a variável não depende das coordenadas no espaço, ou seja, valores amostrais extraídos de diferentes locações no campo se ajustam a uma única distribuição de probabilidades. Desde que estacionariedade e ergodicidade possam ser assumidas e que se disponha de um conjunto de valores amostrais de determinada VR, é possível descrever a variabilidade espacial dessa VR. Alternativamente, pode-se assumir que apenas os dois primeiros momentos da distribuição estatística da VR são invariáveis no espaço, conduzindo à hipótese de estacionariedade de segunda ordem. Os dois primeiros momentos são: valor esperado (momento de primeira ordem) e variância (momento de segunda ordem). Assim, o valor esperado de uma VR denotada por  $W(x)$ , que assume valores  $w(x)$  em cada ponto do espaço com coordenadas  $x$  (uma, duas ou três dimensões), é independente da posição de  $x$  no

espaço e assume um valor constante,  $m$ . De forma genérica, pode-se escrever:

$$E\{W(x)\} = M(x) \quad (1)$$

onde,  $M(x)$  é tendência ou “drift”, que assume o valor constante,  $m$ , quando  $W(x)$  é estacionária. Nesse estudo,  $W(x)$  é assumida como uma VR bidimensional e isotrópica no plano horizontal.

É usual assumir que apenas os incrementos da variável regionalizada são estacionários, conduzindo a hipótese de estacionariedade intrínseca. Estacionariedade de segunda ordem implica em estacionariedade intrínseca, mas o contrário não é verdadeiro. Se a estacionariedade intrínseca é assumida, existe a variância da diferença da VR entre dois pontos no espaço,  $x_1$  e  $x_2$ ,  $[W(x+h) - W(x)]$ , que é função apenas de  $h$ , a distância de separação entre esses pontos e, portanto, não depende das locações  $x_1$  e  $x_2$ , independentemente. Se essa variância existe, pode-se então definir a função semi-variograma como:

$$\gamma(h) = \frac{1}{2}\{\text{var}[W(x)-W(x+h)]\} \quad (2)$$

A função variograma, denominação usual do semi-variograma, que relaciona a metade da variância (semi-variância) das diferenças entre valores observados de  $W(x)$  com a distância de separação desses pontos,  $h$ , define a estrutura de correlação, ou modelo de variabilidade espacial, dessa VR.

A variância, o variograma e a covariância constituem os momentos de segunda ordem de interesse para a geoestatística. Assumindo estacionariedade de segunda ordem, a função variograma,  $\gamma(h)$ , e a função covariância,  $C(h)$ , se relacionam através de:

$$\gamma(h) = C(0)-C(h) \quad (3)$$

Essa relação indica que assumindo-se estacionariedade de segunda ordem, a covariância e o variograma são ferramentas equivalentes para a caracterização da estrutura de correlação da VR (Journel e Huijbregts, 1978).

Nos casos em que a VR possui uma diferença sistemática entre dois valores pontuais, ou seja, uma *tendência* é identificada, o processo é não-estacionário e então a função variograma depende das posições dos pontos  $x_1$  e  $x_2$ , além da distância  $h$ , entre eles.

## Estimativa de variograma experimental e parametrização

A função variograma calculada a partir de valores amostrais de determinada VR é conhecida como variograma experimental ou não-paramétrico. Existem diversos procedimentos para o cálculo do variograma experimental. A escolha do estimador para o variograma experimental depende das características dos dados. O estimador clássico (Matheron, 1963) para o cálculo do variograma requer que os dados amostrais sejam normalmente distribuídos. Cressie e Hawkins (1980) propuseram o estimador "robusto" que é compatível com a presença de um certo grau de "contaminação" nos dados amostrais, ou seja, desvios em relação à distribuição. O estimador robusto é escrito como:

$$\bar{\gamma}(h) = \frac{1}{2[0.457 + 0.494/n(h)]} \left[ \frac{1}{n(h)} \sum_{i=1}^{n(h)} (|w(x_i) - w(x_i + h)|)^{1/2} \right]^4 \quad (4)$$

onde  $n(h)$  é o número de pares de pontos de amostragem separados por  $h$ .

O ajuste de um modelo paramétrico ao variograma experimental é necessário à composição do modelo geoestatístico de uma VR. A escolha primária dos modelos teóricos a serem testados é usualmente condicionada a questões de simplicidade e de desempenho em aplicações anteriores. Os modelos mais frequentemente empregados em estudos geoestatísticos são o exponencial, o esférico e o gaussiano. Qualquer que seja o modelo, os parâmetros a ajustar são aqueles relacionados à caracterização do variograma: o efeito pepita, a soleira, e o alcance. Esses parâmetros definem o modelo geoestatístico através da variância e escala de correlação da VR. As expressões matemáticas dos modelos mais utilizados são:

### 1) Modelo de Gauss

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_S \left[ 1 - \exp\left(-\frac{h^2}{a_g^2}\right) \right] & h \neq 0 \end{cases} \quad (5)$$

### 2) Modelo exponencial

$$\gamma(h, C_0, C_S, a_e) = \begin{cases} 0 & h = 0 \\ C_0 + C_S [1 - \exp(-\frac{|h|}{a_e})] & h \neq 0 \end{cases} \quad (6)$$

### 3) Modelo esférico

$$\gamma(h, C_0, C_S, a_s) = \begin{cases} 0 & h = 0 \\ C_0 + C_S \left[ (3/2) \left( \frac{|h|}{a_s} \right) - (1/2) \left( \frac{|h|}{a_s} \right)^3 \right] & 0 < |h| \leq a_s \\ C_0 + C_S & a_s < |h| \end{cases} \quad (7)$$

onde, para todos os modelos,  $C_0$  representa o efeito pepita, ou a variância correspondente a uma distância de separação inferior à menor distância de amostragem, e  $(C_0+C_S)$  representa a soleira ("sill") do semi-variograma, ou a máxima variância de campo; o parâmetro  $a_g$ , no modelo de Gauss, está relacionado com o alcance do variograma  $a$ , ou à máxima distância de correlação, através de  $a = (3)^{1/2} \cdot a_g$ ; o modelo exponencial atinge o comprimento de correlação apenas assimptoticamente e um valor efetivo para o alcance é determinado como  $a = 3a_e$ ; no modelo esférico,  $a_s$  corresponde ao próprio alcance.

Diversos são os métodos de ajuste do variograma experimental aos modelos paramétricos selecionados para determinação dos parâmetros que definem a estrutura de correlação da VR. Para processos estacionários, ou estatisticamente homogêneos, destaca-se o método dos Mínimos Quadrados Ponderados (MQP). O critério de minimização do MQP proposto por Cressie (1985) considera os termos da matriz de covariância, ou equivalentemente valores assumidos pela função variograma, e é escrito como:

$$\sum_{j=1}^K \frac{|n(h(j))|}{\{\gamma(h(j); \theta)\}^2} [\gamma(h(j)) - \gamma(h(j); \theta)]^2 \quad (8)$$

onde  $\theta$  representa o conjunto de parâmetros a estimar pelo ajuste do variograma teórico  $\gamma$ , ao variograma experimental  $\bar{\gamma}$ ;  $n(h)$  é o número de pares de observações de cada classe de distância de separação,  $h(1), h(2), \dots, h(K)$ ;  $K$  é o número de classes de distância.

A minimização da Equação (7) para obter  $\theta$  tem as seguintes vantagens: 1) as classes de distância próximas à origem do variograma recebem maior peso, o que garante um melhor ajuste nessa região, fator importante para a interpolação através

do kriging; 2) quanto maior o número de pares de dados,  $n(h)$ , em um intervalo de distância  $h$ , maior o peso da diferença quadrática correspondente no ajuste total. Os procedimentos do tipo “mínimos quadrados” em geoestatística têm a vantagem, em relação a outros métodos, de requererem o menor número de hipóteses em relação à distribuição de probabilidade dos dados amostrais. Para a aplicação do método dos mínimos quadrados ponderados não é necessário que os dados estejam em perfeita conformidade com a distribuição normal, sendo permitido um certo grau de “contaminação”.

A forma do variograma experimental pode ser o primeiro elemento para análise da presença de uma tendência no processo. Quando não existe uma definição clara do alcance do variograma admite-se que a VR é não-estacionária, ou seja, existe uma tendência no seu valor esperado ou média. No entanto, Russo e Jury (1987) destacam que nem sempre a presença de uma tendência pode ser detectada pela forma do variograma.

Em casos de VR não estacionárias uma possibilidade para a estimativa do variograma experimental e ajuste do modelo paramétrico é a consideração dos resíduos, através da remoção da tendência espacial. Contudo, diversos autores afirmam que o variograma dos resíduos é uma estimativa tendenciosa do variograma real, devido a interdependência entre a função da tendência filtrada e a função variograma.

No caso de processos não-estacionários, dentre as técnicas de ajuste do variograma experimental a modelos paramétricos destaca-se o método iterativo dos mínimos quadrados dos incrementos genéricos (IGLS, ou “Iteractive Generalised Least Squares”) (Neuman e Jacobson, 1984). Os incrementos genéricos são obtidos filtrando a tendência espacial dos dados amostrais. O IGLS é um procedimento de análise de tendência, e estima simultaneamente, através de procedimento iterativo, os parâmetros do variograma e os parâmetros da função de tendência, desde que as ambas relações funcionais tenham sido previamente definidas.

A função de tendência espacial da VR  $W(x)$  pode ser representada por combinações lineares de funções polinomiais relacionadas às coordenadas espaciais  $f(x)$ , como:

$$M(x) = \sum_{i=0}^{k-1} f_i(x) \cdot \beta_i \quad (9)$$

sendo  $\beta_i$  os parâmetros da tendência e com:

$$k = \frac{(p+1)(p+2)}{2} \quad (10)$$

onde  $p$  é a ordem do polinômio, e  $f_0(x) = 1$ .

O procedimento IGLS para a estimativa dos parâmetros do variograma e da tendência consiste em:

- Assumir  $p = 1$  na Equação (9) e calcular a tendência através do método dos mínimos quadrados (admitindo que as observações não estão correlacionadas espacialmente) utilizando a seguinte expressão escrita em notação matricial:

$$\bar{\beta} = (X^T X)^{-1} X^T W \quad (11)$$

onde  $W$  é o vetor de observações de dimensão ( $N \times 1$ ),  $X$  é matriz de coordenadas espaciais de dimensões ( $k \times N$ ) utilizada em regressão linear padrão, e  $\bar{\beta}$  é o vetor de dimensões ( $k \times 1$ ) dos coeficientes da função de tendência espacial.

- Calcular os resíduos ( $W - X^T \cdot \bar{\beta}$ ) e estimar o variograma experimental. Se, ao proceder a análise visual do variograma, uma tendência é observada sugerindo a existência de um processo não estacionário, então  $p$  é incrementado e o procedimento é repetido a partir do passo 1. Caso contrário executa-se o passo seguinte.
- Ajustar uma função paramétrica ao variograma dos resíduos utilizando um procedimento qualquer para VR estacionária. Com o modelo ajustado, calcula-se a matriz de covariância  $V$ ;
- Estimar os parâmetros da função de tendência levando em consideração a correlação espacial e utilizando o método dos mínimos quadrados ponderados através de  $\bar{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} W$  onde  $V$  é a matriz de covariância de dimensões ( $N \times N$ ).
- Recalcular os resíduos com os novos parâmetros  $\bar{\beta}$  e repetir o procedimento a partir do passo 3 até que a convergência seja alcançada, ou seja, os parâmetros da tendência e do variograma não se alterem substancialmente entre duas iterações consecutivas.

Cressie (1993) afirma que esse procedimento produz estimativas tendenciosas das ordenadas do variograma, mas Clarke (1994) e Russo e Jury (1987) sugerem que essa tendenciosidade é

pequena para intervalos de distância ( $h$ ) próximos à origem do variograma. Esses intervalos de distância próximos à origem do variograma são os de maior influência nos cálculos do kriging.

## Validação e seleção

Uma vez determinados os parâmetros do modelo geoestatístico, ou estrutural, é necessário validar o modelo de variograma ajustado, ou seja, investigar se esse modelo é adequado e consistente com o conjunto de dados amostrais. Dois métodos vêm sendo utilizados para esse propósito: o método da validação cruzada e o método dos resíduos não-correlacionados.

O método dos resíduos não-correlacionados realiza análise estatística no vetor dos resíduos (dados originais filtrando a tendência espacial) transformado em quantidades não-correlacionadas. Os resíduos da VR  $W(x)$ ,  $z(x) = w(x) - M(x)$ , são transformados em quantidades não relacionadas,  $u(x)$ , através da expressão seguinte escrita em notação matricial:

$$u = C^{-1}z \quad (12)$$

onde a matriz  $C$  é obtida através da matriz de covariância  $V$ , a partir da seguinte decomposição:

$$CC^T = V \quad (13)$$

Para o modelo do variograma ser válido, os elementos do vetor transformado,  $u(x)$ , devem satisfazer a condição de serem variáveis normalizadas independentes, pertencentes a distribuição Normal padrão,  $N(0,1)$ . Com um nível de confiança de 95%, normalidade e independência são satisfeitas desde que as seguintes condições, respectivamente, sejam alcançadas:

$$-2 \leq u(x_i) \leq 2, \text{ para } i = 1, \dots, N \quad (14)$$

e

$$-2 \leq u(x_i) \cdot u(x_j) \leq 2, \text{ para } i, j = 1, \dots, N \text{ e } i \neq j \quad (15)$$

O teste de qui-quadrado pode ser aplicado para confirmar a hipótese de  $u(x)$  seguir uma distribuição normal-padrão,  $N(0,1)$ .

O método da validação cruzada ou “jacknifing” é baseado na técnica de interpolação de kriging. O procedimento consiste em produzir uma estimativa  $w^*$  em um ponto de observação  $x$ , através da técnica de interpolação de kriging, utilizando o restante das  $(N-1)$  observações nas demais loca-

ções. O processo é repetido para cada ponto de observação no espaço. O vetor do erros produzidos por essas estimativas é calculado na sua forma reduzida, ou seja, considerando a variância do kriging, como:

$$RE(x_i) = \frac{[w^*(x_i) - w(x_i)]}{(\text{var}[w^*(x_i) - w(x_i)])^{1/2}} \quad (16)$$

Para o modelo testado ser consistente com o conjunto de observações,  $RE(x)$  deve ser normalmente distribuído, com média zero, implicando na ausência de erros sistemáticos de superestimação ou subestimação, e variância unitária, o que garante a consistência entre a variância do kriging e a variância do erro correspondente  $[w^*(x_i) - w(x_i)]$ . Essas duas condições são satisfeitas desde que:

$$ME = \frac{1}{N} \sum_{i=1}^N RE(x_i) \approx 0 \quad (17)$$

e

$$MRE = \left[ \frac{1}{N} \sum_{i=1}^N [RE(x_i)]^2 \right]^{1/2} \approx 1 \quad (18)$$

onde ME corresponde à média dos erros reduzidos e MRE corresponde à média quadrática dos erros reduzidos. Adicionalmente a essas duas condições, uma aferição do ajuste global à distribuição normal deve ser realizada, para garantir que o modelo de variograma e os parâmetros testados são uma descrição precisa da estrutura de correlação da VR.

O teste pelos métodos descritos pode validar vários modelos de variograma para um mesmo conjunto de dados representando uma VR. A seleção do modelo mais apropriado, dentre uma série de modelos classificados como adequados e consistentes com o conjunto de dados amostrais, deve ser baseada em um critério de discriminação. O AIC (“Akaike Information Criterion”) pode ser utilizado para esse fim e é escrito como:

$$AIC = 2(L + K_i) \quad (19)$$

onde  $L$  é logaritmo negativo da verossimilhança para um modelo particular, e  $K_i$  é o número de parâmetros independentes ajustados para ambas as funções do variograma e da tendência espacial. O modelo com melhor ajuste corresponde ao de AIC mínimo. A expressão para o cálculo de  $L$  quando os dados são valores transformados em logaritmos é dada por Hoeksema e Kitanidis (1985).

$$L = \frac{N}{2} \ln(2\pi) + \frac{1}{2} \ln(\det(V)) + \frac{1}{2} (W - E(W))^T V^{-1} (W - E(W)) + \sum_{i=1}^N w_i \quad (20)$$

## ÁREA DE ESTUDO E DADOS UTILIZADOS

O domínio de estudo corresponde à parte irrigável da Fazenda Nossa Senhora do Rosário, situada no município de Pesqueira, na região agreste do Estado de Pernambuco. A área de estudo corresponde à cerca de 60 ha. A área está localizada sobre um aquífero aluvial com profundidades do lençol que variam espacialmente e sazonalmente entre 2,0 e 6,0 m.

Foram coletadas amostras de solo em 55 locações na área, considerando três profundidades distintas: 30, 90 e 250 cm. Os pontos de amostragem foram locados aleatoriamente na área. As distâncias mínima e máxima entre os pontos de observação correspondem a 5 e 1400 metros, respectivamente.

As amostras foram analisadas em laboratório para determinação da textura (conteúdo de areia, silte e argila). Os dados de textura foram utilizados para cálculo da condutividade hidráulica saturada ( $K_{sat}$ ) através de metodologia indireta conhecida como Funções de Transferência Hidropedológica (FTH). Um total de 42 amostras, incluindo as três profundidades, foram utilizadas para testar a FTH para o cálculo da condutividade hidráulica saturada. Medidas de condutividade hidráulica saturada em laboratório utilizando amostras indeformadas foram procedidas utilizando permeímetro de parede flexível e permeímetro de parede rígida. As amostras foram previamente saturadas com dióxido de carbono e então submetidas ao ensaio utilizando água de aerada. Montenegro (1997a) e Montenegro (1997b) detalham os procedimentos de coleta de amostras em campo e ensaios em laboratório. Foram testadas as FTH desenvolvidas por Rawls e Brakensiek (1989) e por Saxton et al. (1986), com duas combinações diferentes para a estimativa da porosidade. As funções foram testadas estatisticamente quanto à significância da diferença entre media e variância em relação ao conjunto de dados de laboratório. Para os testes foram considerados os logaritmos da condutividade hidráulica saturada. Ainda, foram calculados os coeficientes de correlação entre os valores de laboratório e os avaliados por cada uma das funções testadas. A descrição do processo de teste e seleção da FTH de melhor adequação é

descrito por Montenegro e Montenegro (1997). Para a caracterização da variabilidade espacial de uma VR através da geoestatística é fundamental a definição dos dois primeiros momentos da distribuição de probabilidades. Os testes estatísticos realizados (significância da diferença entre média e variância de duas populações representadas por amostras de estimativas de laboratório e estimativa de FTH) provaram que esses dois momentos são preservados quando se considera os valores de  $\log(K_{sat})$  avaliados pela FTH selecionada para representar os valores medidos experimentalmente. A vantagem dessa metodologia indireta é a simplificação da tarefa de obtenção de dados para a caracterização de uma estrutura de variabilidade espacial. Ainda, deve-se observar que, uma vez que a abordagem estocástica espacial baseada em *kriging* e *geração de cenários condicionais* é capaz de levar em consideração ambas a variabilidade espacial e a incerteza nos valores dos parâmetros (Delhomme, 1979), essa metodologia é potencialmente aplicável.

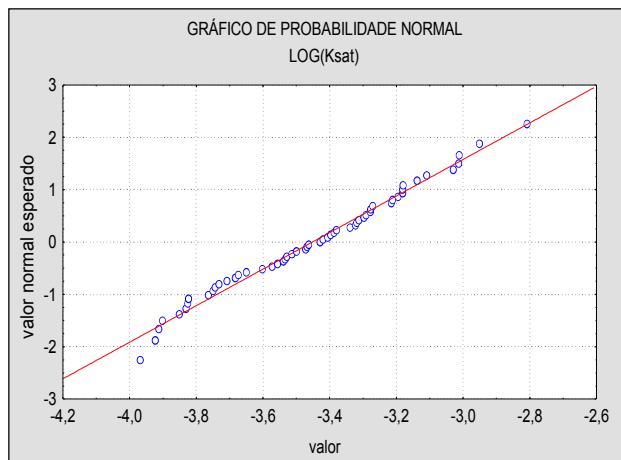
## ESTUDO DE VARIABILIDADE ESPACIAL: RESULTADOS E DISCUSSÃO

### Análise estatística clássica

Como hipótese da modelagem estocástica pretendida,  $K_{sat}$  em cada locação é avaliada por um valor médio, tomando-se as três profundidades amostradas, representando uma coluna de solo homogênea (Montenegro, 1997a). Foi assumido, como ponto de partida para a análise, que o comportamento estatístico da condutividade hidráulica é melhor representado por uma distribuição log-normal. Portanto, o valor médio associado à distribuição da condutividade hidráulica é dado pela média aritmética dos valores dos logaritmos de  $K_{sat}$ , e que corresponde a média geométrica dos valores de  $K_{sat}$ . A hipótese de normalidade para os valores dos logaritmos de  $K_{sat}$  é testada através de dois critérios: o teste gráfico e a aplicação do teste K-S (Kolmogorov-Smirnov). O teste gráfico pode ser visualizado na Figura 1. Observa-se que a maioria dos pontos são distribuídos ao longo de uma reta, indicando portanto, um razoável ajuste à distribuição Normal.

O teste K-S também foi aplicado ao conjunto de dados para confirmar a normalidade. O teste K-S é do tipo não-paramétrico e é caracterizado por ser menos sensível aos extremos da distribuição do que os testes do tipo paramétrico. A

hipótese nula foi formulada como a de distribuição normal para os valores logaritmos de  $K_{sat}$ . A estatística do teste foi calculada como 0,0617. A hipótese nula de distribuição Normal não foi rejeitada a um nível de confiança de 95%, uma vez que o valor crítico correspondente é de 0,183.



**Figura 1. Teste gráfico da adequação da distribuição normal.**

Uma vez confirmada a hipótese de normalidade, a variável Log ( $K_{sat}$ ) passa a ser representada por uma distribuição Normal com média, -3,45, e desvio-padrão 0,278. Porém, uma vez que a estrutura de correlação não foi levada em consideração, tratando-se os valores de log ( $K_{sat}$ ) como independentes e não-correlacionados, o resultado dessa análise pode ser tendencioso, não refletindo a realidade.

## Caracterização da estrutura de correlação: análise geoestatística

Para o cálculo dos variogramas experimentais foi utilizado o estimador robusto (Equação 4). Apesar de confirmada a hipótese de normalidade, é provável a existência de "contaminação" no conjunto de dados, ou seja, pequenos desvios em relação à distribuição normal. Isso pode ser constatado através da observação dos extremos da distribuição no gráfico da Figura 1.

O variograma experimental foi calculado considerando um arranjo dos dados tal que  $n(h) \geq 30$  (Equação 4) para cada classe de distância, de acordo com o sugerido por Journel e Huijbregts (1978).

O variograma experimental (Figura 2) apresenta flutuações, caracterizando um certo com-

portamento periódico. Essas irregularidades podem ser devidas ao fato de  $h$  corresponder a um valor médio representando um intervalo de classes de distância, uma vez que as distâncias de amostragem não foram regulares. Essa é uma das possíveis causas de irregularidades em variogramas apontadas por Russo (1984). Outra causa provável para as irregularidades do variograma pode ser a presença de valores extremos ("outliers") no conjunto de dados de log ( $K_{sat}$ ), devido por exemplo à presença de lentes de argila. A não consideração desses valores poderia resultar em uma melhoria da forma do variograma, mas em contrapartida poderia se obter uma estimativa tendenciosa. Esse fato é ressaltado por Woodbury e Sudicky (1991) em estudo de caracterização geoestatística de aquífero. Destaca-se o fato de que o variograma dos log ( $K_{sat}$ ) na profundidade 30 cm (não apresentado) evidenciou o mesmo comportamento periódico do variograma dos valores médios, embora as variâncias calculadas tenham sido geralmente maiores no caso do variograma de 30 cm em relação ao variograma dos valores médios, assim como o comprimento de correlação. Esse comportamento periódico não foi observado nos variogramas correspondentes às profundidades de 90 e 250 cm. Esse fato pode ser explicado considerando o efeito do uso do solo traduzido na observação de duas categorias de solos distintas: os solos cultivados e os não cultivados. Esse efeito evidentemente é mais notável nas camadas mais superficiais do solo, como a correspondente à profundidade de amostragem de 30 cm. Essa poderia ser indicada como uma das causas do comportamento periódico do variograma do log ( $K_{sat}$ ) médio no perfil, sendo este fortemente influenciado pelo comportamento da mesma variável a 30 cm de profundidade. Isso caracterizaria a chamada variabilidade extrínseca. De acordo com Rao e Wagenet (1985), a variabilidade extrínseca pode ser originada, em parte, em decorrência de práticas de produção agrícola, enquanto que a variabilidade intrínseca é aquela relacionada a variações naturais nos solos. Journel e Huijbregts (1978) classificam esse tipo de ocorrência no variograma experimental de componente pseudo-periódica e exemplificam sua ocorrência num caso em que se observa a existência de dois tipos bem diferenciados de mineralização, presentes em uma área de depósito de minérios. Mais investigações à cerca da causa dessas irregularidades no variograma experimental seriam necessárias para o caso em estudo.

Para a parametrização do variograma experimental, três modelos foram selecionados a priori como candidatos: esférico, exponencial e gaussiano. Uma tendência ("drift") linear também foi investiga-

da, embora o exame visual do variograma experimental não tenha indicado sua presença. Portanto, para cada um dos modelos testados, dois possíveis casos foram investigados, um com o valor esperado constante e outro com a tendência linear. O ajuste de cada um dos seis modelos paramétricos foi feito através de uma subrotina computacional inserida na rotina de regressão não-linear do sistema SAS (SAS Institute Inc., 1985), baseada no método IGLS. O método dos mínimos quadrados ponderados foi utilizado para o ajuste do variograma das observações, no caso de modelo considerando valor esperado constante, e para o ajuste do variograma dos resíduos (passo 3 do procedimento IGLS descrito) no caso do modelo considerar a tendência espacial. Segundo a orientação de Cressie (1993), o ajuste de modelos paramétricos considera apenas as ordenadas do variograma até a metade do máximo intervalo de distância. Para facilitar a convergência do processo e garantir identificabilidade dos parâmetros do modelo de variograma individualmente, foram estabelecidos limites para o parâmetro relacionado ao comprimento de correlação, segundo a sugestão de Hoeksema e Kitanidis (1985). O limite superior corresponde à máxima distância de separação dos pontos amostrais e o limite inferior é dado pelo maior valor entre a mínima distância de separação e a décima parte da máxima distância de separação. Os parâmetros dos modelos ajustados estão listados na Tabela 1.

Pode ser observado que a escolha do modelo tem influência na obtenção dos valores da variância de campo ( $C_0+C_s$ ) e do comprimento de correlação, descrevendo a estrutura de correlação de  $\log(K_{sat})$  no domínio em estudo. Os coeficientes estimados para a tendência linear são similares para os três modelos considerados. A consideração da correlação espacial, através da geoestatís-

tica, tem influência marcante no cálculo da variância do processo, como pode ser notado comparando o valor de ( $C_0+C_s$ ) ajustado para cada um dos modelos considerados com a variância obtida pela análise estatística clássica. Para todos os modelos ajustados a variância de campo estimada pela geoestatística foi superior ao valor estimado sem considerar a correlação espacial, ou seja, 0,078.

Considerando os seis modelos ajustados ao variograma experimental, a validação foi feita segundo os dois critérios anteriormente descritos: a validação cruzada e a análise dos resíduos não-correlacionados. O aplicativo computacional GEO-EAS (Englund e Sparks, 1992) foi utilizado para os cálculos do procedimento da validação cruzada. Os testes de significância de ME e MRE (t-teste), no procedimento da validação cruzada, na média (t-teste) e variância (teste qui-quadrado) na análise dos resíduos não-correlacionados, foram realizados considerando um nível de confiança de 95%. O teste de Kolmogorov-Smirnof (K-S) foi utilizado para testar a normalidade. Os resultados dos testes para os seis modelos considerados são apresentados na Tabela 2.

O critério de discriminação AIC (Equação 15) foi calculado para todos os modelos. O número de variáveis independentes,  $K_i$ , na expressão de AIC é igual a 6 para os modelos com tendência espacial, e igual a 4 para os modelos de média constante. Os resultados são apresentados na Tabela 2. De acordo com os resultados da Tabela 2, apenas o modelo gaussiano com valor esperado constante foi validado pelos dois procedimentos aplicados. Esse foi portanto o modelo escolhido. A Figura 2 apresenta o variograma experimental e o modelo de gauss ajustado, cujos parâmetros encontram-se listados na Tabela 1.

**Tabela 1. Parâmetros de diferentes modelos de variograma de  $\log(K_{sat})$  médio no perfil de solo.**

Modelo**	Variograma			Tendência*		
	$C_0$	$C_s$	Comprimento de correlação (m)	$\beta_0$	$\beta_1$	$\beta_2$
E+C	0,0376	0,0657	163,77	-3,45	-	-
S+C	0,0453	0,049	177,8	-3,45	-	-
G+C	0,043	0,0623	166,61	-3,45	-	-
E+L	0,023	0,078	145,95	-11,59	$-8,4 \times 10^{-5}$	$9,1 \times 10^{-4}$
S+L	0,036	0,0656	193,59	-11,54	$-8,7 \times 10^{-5}$	$9,1 \times 10^{-4}$
G+L	0,0399	0,06	172,78	-11,38	$-7 \times 10^{-5}$	$8,7 \times 10^{-4}$

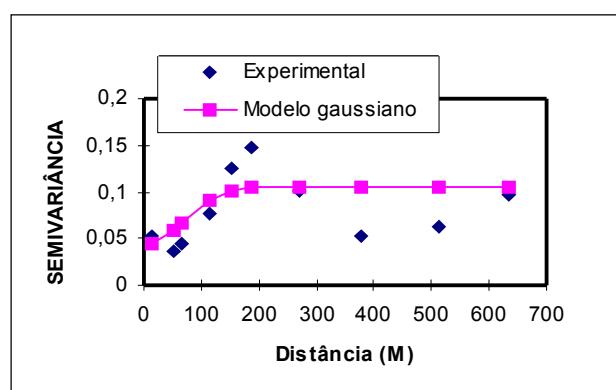
\* A tendência é calculada como  $M(x,y) = \beta_0 + \beta_1 x + \beta_2 y$ ; \*\* E, S e G referem-se aos modelos exponencial, esférico e Gaussiano, respectivamente; C e L referem-se a tendência constante e linear, respectivamente.

**Tabela 2. Parâmetros dos testes de validação e seleção de modelos de variograma.**

Modelo	Validação cruzada					Resíduos não-correlacionados				
	ME	MRE	MSE	correl <sup>‡</sup>	K-S	$\mu$	$\sigma$	K-S	SSR	AIC
E+C	0,007	0,778 <sup>†</sup>	0,214	0,635	0,1094	-0,006	1,042	0,0907	58,73	-360,9*
S+C	0,007	0,962	0,211	0,649	0,1255	-0,0038	1,28	0,1118	88,64 <sup>†</sup>	-339,8*
G+C	0,003	0,843	0,213	0,64	0,1083	0,003	0,843	0,08	71,66	-353,5
E+L	0,011	0,85	0,212	0,646	0,1205	0,052	1,21	0,102	79,28 <sup>†</sup>	-351,7*
S+L	0,013	0,837 <sup>†</sup>	0,208	0,655	0,121	-0,0059	1,182	0,085	75,5 <sup>†</sup>	-356,6**
G+L	0,09	0,865	0,207	0,656	0,1139	-0,0263	1,22	0,0997	80,41 <sup>†</sup>	-353,5*

\* diferença significante a um nível de confiança de 95%; † coeficiente de correlação do ajuste (regressão não-linear);

\*o modelo falhou em um dos testes de validação; \*\*o modelo falhou em ambos testes de validação; valor crítico teste K-S = 0,183; valores listados do teste K-S referem-se às estatísticas calculadas.



**Figura 2. Variogramas, experimental e modelo ajustado, de log( $K_{sat}$ ).**

### Aplicação do modelo de variabilidade da condutividade hidráulica

O processo de validação garante que o modelo teórico reproduz o comportamento da VR, em termos de variabilidade espacial inferido através do variograma experimental. O variograma paramétrico selecionado identifica a estrutura de correlação de  $K_{sat}$  e compõe o modelo de variabilidade espacial do parâmetro. Esse modelo é utilizado para interpolação de  $\log(K_{sat})$  na área em estudo através de kriging (Figura 3a). Foi utilizado o kriging pontual em células quadradas de lado igual a 10 metros, valor considerado adequado para resolução da estrutura de correlação identificada. A interpolação através do kriging também permite calcular a variância associada aos dados interpolados, ou a faixa de erro dos valores estimados. A Figura 4 apresenta mapa de contorno dos desvios-padrão do kriging de  $\log(K_{sat})$ . A variância do kriging em qualquer ponto  $x_i$  do monínio está

relacionada com a precisão da estimativa da VR nesse ponto e pode ser utilizada para definir o intervalo de confiança, 95% por exemplo, das estimativas do kriging. A estrutura espacial da variância do kriging não depende dos valores observados, mas do modelo de variograma e da configuração geométrica dos pontos de observação. A análise da variância do kriging pode ser utilizada tanto para selecionar o modelo de variograma como para propor um refinamento da malha de pontos de observação. Como esperado, os valores dos desvios-padrão são menores nas proximidades dos pontos de observação, apesar de ainda elevados devido à presença da micro-estrutura de correlação, identificada através do efeito pepita no variograma. Os desvios-padrão do kriging são próximos aos desvios-padrão de campo (0,27) em uma grande proporção da área, indicando estimativas não muito precisas. Esse fato se deve provavelmente ao reduzido número de pontos de observação.

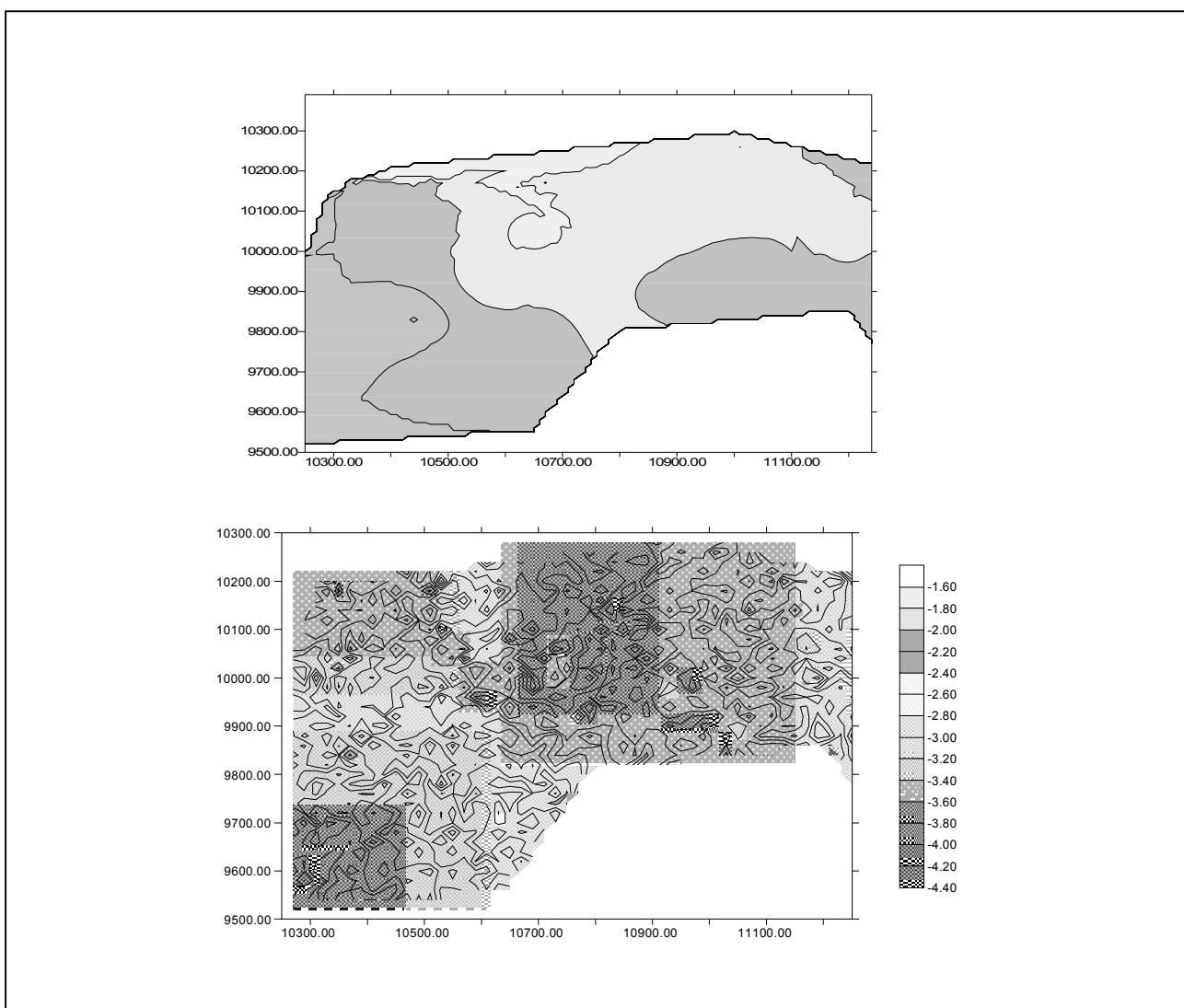
Estudos destacam que os parâmetros dos variogramas ajustados não devem ser considerados como determinísticos, mas a eles deve ser associada uma componente de incerteza (Woodbury e Sudicky, 1991). Todini e Ferraresi (1996), demonstraram que essa incerteza tem uma influência marcante nas estimativas do kriging e nas variâncias associadas. Essa abordagem não foi levada em consideração nesse estudo.

O processo de interpolação apenas calcula os valores das VR em pontos diferentes aos de observação, sem levar em consideração a incerteza decorrente de erros de estimativa e da própria variabilidade espacial. Essa incerteza, que é a base para a modelagem espacial estocástica, é caracterizada a partir da geração de um número finito de possíveis cenários de distribuição espacial da VR. Essa geração pode ser do tipo *condicional*

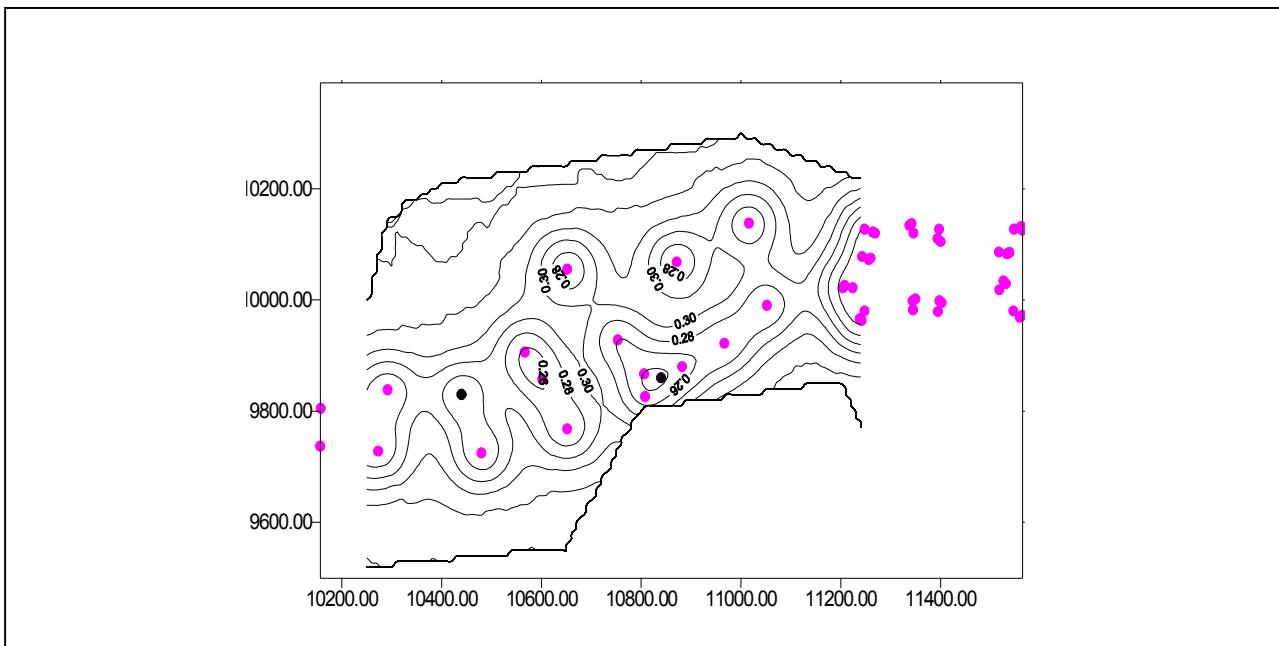
ou *não-condicional*. Em ambos os casos o modelo de variabilidade espacial é utilizado, mas apenas nas chamadas simulações condicionais os valores observados são reproduzidos nos pontos de medida. Para fins de ilustração, apresenta-se na Figura 3b o mapa de contorno de uma simulação não-condicional, pertencente a uma série de 50 simulações de  $\log(K_{\text{sat}})$ . Comparando-se a Figura 3a com a Figura 3b observa-se que parte da variabilidade espacial da VR é suavizada pelo processo de interpolação, enquanto que a simulação condicional resgata essa variabilidade. A série de simulações, ou cenários, gerados é a base para a modelagem estocástica espacial de escoamento em solos não saturados, onde se considera a incerteza associada à variabilidade espacial e a erros de estimação dos parâmetros.

## CONSIDERAÇÕES FINAIS

Demonstrou-se o uso da geoestatística na caracterização da variabilidade espacial de parâmetro hidráulico em uma área. Embora o modelo de variograma escolhido para descrição da variabilidade espacial do  $\log(K_{\text{sat}})$  médio no perfil do solo tenha sido aprovado em testes de adequação e precisão, análises adicionais podem ser procedidas, considerando:



**Figura 3a. Mapa de contorno do kriging de  $\log(K_{\text{sat}})$  (cm/s) na área. 3b. Possível cenário de variabilidade de  $\log(K_{\text{sat}})$  na área.**



**Figura 4. Mapa de contorno dos desvios-padrão do kriging (pontos representam os locais de amostragem).**

- outros modelos de variograma como candidatos;
- variogramas direcionais, ou seja, anisotropia;
- teste de outras funções de tendência que não a linear.

Além disso, investigações a respeito da influência do esquema de amostragem (número e distribuição dos pontos de observação) utilizado na determinação do variograma experimental, também podem ser incluídas em estudos futuros.

O interesse principal da caracterização da variabilidade espacial realizada nesse estudo é fornecer elementos para a modelagem estocástica espacial de fluxo de água e transporte de solutos no solo não-saturado. Dessa forma, merece consideração a avaliação funcional do resultado dessa análise de variabilidade espacial. Ou seja, a investigação da influência de aspectos selecionados da análise estrutural, como os anteriormente mencionados, em resultados específicos da modelagem, como por exemplo, a distribuição de frequência da umidade média na zona radicular da área estudada, ao final de um ciclo de cultivo. Aspectos aparentemente relevantes, como irregularidades no variograma experimental, podem ser demonstrados como de importância secundária através da análise funcional.

O estudo também demonstra os efeitos da suavização da variabilidade espacial, produzidos pela interpolação através do kriging. Em análises de transporte de solutos, onde a escala de heterogeneidade é relevante, é recomendável a geração de cenários condicionais da propriedade hidráulica, conduzindo assim a uma abordagem estocástica.

Finalmente, o estudo apresentado focaliza a questão da análise de variabilidade espacial de valores de  $K_{sat}$ , obtidos indiretamente a partir de funções de transferência hidropedológica. Embora a validação dessas funções não seja o objeto desse trabalho, é importante destacar a validade do seu uso em caracterização de variabilidade espacial, pelas razões que se seguem:

- a condutividade hidráulica saturada é um parâmetro físico de determinação não muito fácil. Os valores medidos em campo têm um nível de incerteza inerente aos procedimentos de medida;
- na análise estocástica, os parâmetros são considerados valores incertos e uma grande quantidade de estimativas pontuais é requerida para avaliação da estrutura de correlação espacial, tornando o processo dispendioso e demorado. A modelagem determinística, apesar de imprecisa, tem sido aplicada frequentemente por falta de dados suficientes para a caracterização da estrutura espacial do parâmetro;

- uma alternativa para avaliação da estrutura de correlação entre parâmetros de difícil obtenção é o uso de co-kriging, que utiliza correlações que, por ventura existam entre esses parâmetros e outros de mais fácil determinação, e que portanto, podem ter sua estrutura espacial inferida mais precisamente (Vauclin et al., 1983; Vaughan et al., 1995). No caso de  $K_{sat}$ , sua estrutura de correlação pode ser inferida a partir da variabilidade espacial de valores correlacionados, como quantidades relacionadas à textura. Essa constitui metodologia indireta de emprego usual para superar os problemas destacados no parágrafo anterior. Pe- lo seu potencial de representação baseado também em correlações, as FTH também devem ser consideradas como metodologia indireta de uso potencial;
- outras fontes de incerteza em geral não incorporadas à análises de variabilidade espacial, como por exemplo o caráter estocástico dos parâmetros do variograma ajustado a partir dos dados observados, podem ter uma influência muito mais significativa nos resultados da modelagem do que os erros associados às estimativas da FTH;
- o impacto do uso de estimativas indiretas na modelagem pode ser minimizado se forem considerados cenários de variabilidade espacial do tipo *não-condicional*, onde apenas a estrutura de correlação espacial, inferida indiretamente pelas FTH, é utilizada para representar a incerteza advinda da variabilidade, sem levar em consideração os valores pontuais;
- o impacto do uso dessas FTH na modelagem estocástica de fluxo e transporte de solutos também deve ser avaliado. No entanto, isso se constitui em objeto de estudo mais específico para a validação de FTH.

## AGRADECIMENTOS

Os autores agradecem ao CNPq, à Universidade de Newcastle (Inglaterra), à UFPE, à UFRPE e à EMATER-PE pelo apoio à pesquisa desenvolvida em programa de doutoramento.

## REFERÊNCIAS

- CAMBARDELLA, C. A.; T. B. MOORMAN; J. M. NOVAK; T. B. PARKIN; D. L. KARLEN; R. F. TURCO, & A. E. KONOPKA (1994). Field-scale variability of soil properties in Central Iowa Soils. *Soil Sci. Soc. Am. J.*, 58: 1501-1511.
- CLARKE, R. T. (1994). *Statistical Modelling in Hydrology*. John Wiley & Sons Ltd. p412 England.
- CRESSIE, N. D. & HAWKINS, D. (1980). Robust estimation of the variogram, I. *Journal of the International Association for Mathematical Geology*, 12: 115-125.
- CRESSIE, N. D. (1985). Fitting models by weighted least squares. *J. Math. Geology*, Vol. 17, No. 4: 563-586.
- CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley. p900. New York.
- ENGLUND, E. & A. SPARKS (1992). GEO-EAS. *Geostatistical Environmental Assessment Software*. International Ground Water Modeling Center. Las Vegas, Nevada.
- DELHOMME, J. P. (1979). Spatial variability and uncertainty in ground water flow parameters: A geostatistical approach. *Water Resources Research*, 15: 269-280.
- HOEKSEMA, R. J. & P. K. KITANIDIS (1985). Analysis of the spatial structure of properties of selected aquifers. *Water Resources Research*, Vol. 21: 563-572.
- JOURNEL, A. G. & CH. J. HUIJBREGTS (1978). *Mining Geostatistics*. Academic Press. 599p. England.
- MATHERON, G. (1963). *Principles of Geostatistics. Economic Geology*, Vol. 58: 1246-1266.
- MONTENEGRO, S. M. G. L. (1997a). *Stochastic analysis of salt accumulation in heterogeneous irrigated soils: a case study in Northeast Brazil*. Tese de Ph. D. Civil Engineering Department, University of Newcastle. Inglaterra.
- MONTENEGRO, A. A. A. (1997b). *Stochastic Hydrogeological Modelling of Aquifer Salinization from Small Scale Agriculture in Northeast Brazil*, Tese de Ph. D., University of Newcastle Upon Tyne. Inglaterra.
- MONTENEGRO, S. M. G. L. & MONTENEGRO, A. A. A. (1997). Uso de funções de transferência hidropedológica em uma área irrigada do Estado de Pernambuco visando aplicação em modelagem espacial estocástica. *Anais do XII Simpósio Brasileiro em Recursos Hídricos*. Vitória, ES. CD.
- NEUMAN, S. P. & E. A. JACOBSON (1984). Analysis of nonintrinsic spatial variability by

- residual kriging with application to regional groundwater levels. *Journal Int. Assoc. Math. Geol.*, 16: 499-521.
- NIELSEN, D. R.; J. W. BIGGAR & K. T. ERH. (1973). Spatial variability of field measured soil-water properties. *Hilgardia*, 42 (7): 215-259.
- RAO, P. S. C. & R. J. WAGENET. (1985). Spatial variability of pesticides in field soils: Methods for data analysis and consequences. *Weed Sci.*, 33 (Suppl. 2): 18-24.
- RAWLS, W. J. & D. L. BRAKENSIEK. (1989). Estimation of soil water retention and hydraulic properties. In: H. J. MOREL-SEYTOUX (ed.), *Unsaturated Flow in Hydrologic Modeling Theory and Practice*, 275-300, Kluwer Academic Publisher.
- RUSSO, D. & E. BRESLER (1981). Soil hydraulic properties as stochastic processes, I. An analysis of field spatial variability. *Soil Sci. Soc. Am. J.*, 46: 20-26.
- RUSSO, D. (1984). Design of an optimal sampling network for estimating the variogram. *Soil Sci. Soc. Am. J.*, Vol. 48: 708-716.
- RUSSO, D. & W. A. JURY (1987). A theoretical study of the estimation of the correlation scale in spatially variable fields, 2. Nostationary fields. *Water Resources Research*, Vol. 23, No. 7: 1269-1279.
- SAS, INSTITUTE INC. (1985). *The NLIN procedure in SAS User's Guide: Statistics*, Version 5. Chap. 5: 575-606.
- SAXTON, D. E., W. J. RAWLS, J. S. RESEMBERGER, & R. I. PAPENDICK (1986). Estimating generalized soil-water characteristics from texture. *Soil Sci. Am. Journal*, 50:1031-1036.
- TODINI, E. & M. FERRARESI (1996). Influence of parameter estimation uncertainty in kriging. *Journal of Hydrology*, 175: 555-566.
- VAUCLIN, M., S. R. VIEIRA, G. VACHAUD, & D. R. NIELSEN (1983). The use of cokriging with limited field soil observations. *Soil Sci. Soc. Am. J.*, 47: 175-184.
- VAUGHAN, P. J., S. M. LESCH, D. L. CORWIN, & D. G. CONE (1995). Water content effect on soil salinity prediction: a geostatistical study using Cokriging. *Soil Sci. Soc. Am. J.*, 59: 1146-1156.
- WOODBURY, A. D. & E. A. SUDICKY (1991). The geostatistical characteristics of the Borden Aquifer. *Water Resources Research*, Vol. 27, No. 4: 533-546.

## ***Spatial Variability of Hydraulic Parameters in Alluvial Soils in the State of Pernambuco***

### **ABSTRACT**

*The spatial variability of saturated hydraulic conductivity is analysed using both statistical and geostatistical methods. The importance of the latter has been highlighted to infer the parameter correlation structure. Data from an irrigated area in Pernambuco State was used. The parameter spatial variability model is comprised by the identified correlation structure, which is the basis to interpolate the parameter outside the measurement points, through kriging.*