

MÉTODO CONVENCIONAL E “BOOTSTRAP” PARA ESTIMAR O NÚMERO DE OBSERVAÇÕES NA DETERMINAÇÃO DOS PARÂMETROS DA FUNÇÃO $K(\theta)$ ⁽¹⁾

J. F. MELO FILHO⁽²⁾, P. L. LIBARDI^(3, 4),
Q. de JONG van LIER^(3, 4) & J. E. CORRENTE⁽³⁾

RESUMO

A variabilidade natural do solo dificulta a obtenção de valores que representem adequadamente as propriedades do solo em determinada área. O conhecimento do número mínimo de observações, que devem ser realizadas para representar, com um erro aceitável, uma propriedade ou característica do solo, é fundamental para que os resultados experimentais possam ser aplicados com segurança. No presente trabalho, comparou-se o método convencional (teste-t) com o método “bootstrap”, com vistas em estimar o número de observações necessárias para calcular os parâmetros que caracterizam a relação entre a condutividade hidráulica e o teor de água do solo, determinada pelo método do perfil instantâneo. Realizou-se um experimento de drenagem num Latossolo Vermelho-Amarelo em Piracicaba (SP), numa parcela experimental com 45 pontos de observação distanciados de 1 m entre si. Observaram-se a umidade (com TDR) e o potencial mátrico (com tensiômetros) durante três semanas de redistribuição da água. Após processamento dos dados, o conjunto de valores mostrou uma distribuição não-normal, fazendo-se necessária a eliminação de “outliers” para a aplicação do método convencional, normalizando a distribuição. Assim, o uso do método tradicional só é recomendado após a confirmação da pertinência da eliminação dos “outliers”. Ambos os métodos de análise requerem grande número de repetições, reafirmando que determinações da função condutividade hidráulica com poucas repetições não podem ser extrapoladas para áreas maiores.

Termos de indexação: condutividade hidráulica, variabilidade, “bootstrap”

⁽¹⁾ Parte da Tese de Doutorado do primeiro autor, apresentada à Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ/USP. Recebido para publicação em dezembro de 2001 e aprovado em setembro de 2002.

⁽²⁾ Professor da Escola de Agronomia, Universidade Federal da Bahia – UFBA. CEP 44380-000 Cruz das Almas (BA). E-mail: jfmelo@esalq.usp.br

⁽³⁾ Professor do Departamento de Ciências Exatas, Escola Superior de Agricultura Luiz de Queiroz – ESALQ/USP. Caixa Postal 9, CEP 13418-970 Piracicaba (SP). E-mail: pllibard@esalq.usp.br

⁽⁴⁾ Bolsista do CNPq.

SUMMARY: *CONVENTIONAL AND BOOTSTRAP METHODS FOR THE ESTIMATION OF THE NUMBER OF OBSERVATIONS THAT DETERMINE $K(\theta)$ FUNCTION PARAMETERS*

Natural variability of soils makes it difficult to obtain values that adequately represent the soil properties within a certain area. The establishment of a minimum number of observations necessary to represent a soil property (with an acceptable error) is fundamental for the reliability of experimental results. In the present paper, the conventional t-test method has been compared with the bootstrap method, in order to estimate the number of observations needed to calculate the parameters that characterize the relation between hydraulic conductivity and soil water content, determined by the instantaneous profile method. A drainage experiment was realized in a red-yellow Oxisol (Hapludox) in Piracicaba, State of São Paulo, Brazil, on an experimental plot with 45 observation points in a straight line with a distance of one meter between points. Soil moisture was measured by TDR and the matric potential by tensiometers during three weeks of water redistribution. After data-processing, the parameters of the $K(\theta)$ function presented a non-normal distribution and outliers had to be eliminated in order to normalize the data set for the t-test application. Therefore, the conventional method should only be used if the elimination of outliers is supposedly reasonable. Both analysis methods indicated the need of a high number of repetitions, reaffirming that hydraulic conductivity functions based on few repetitions should not be extrapolated to larger areas.

Index terms: hydraulic conductivity, variability, bootstrap.

INTRODUÇÃO

O solo é um corpo natural, complexo e heterogêneo, cuja variabilidade dificulta a obtenção, em estudos de campo, de valores médios que representem adequadamente suas propriedades em determinada área, para posterior aplicação na solução de problemas relacionados com a agricultura e com o ambiente. Diante deste fato, em estudos de variabilidade de propriedades do solo, uma pergunta a ser respondida está relacionada com o número mínimo de amostras que deve ser obtido para representar adequadamente, com um erro aceitável, a propriedade em estudo. Segundo McBratney & Webster (1983), a resposta depende da variabilidade da propriedade na área estudada e, atualmente, só pode ser determinada empiricamente, podendo, em alguns casos, alcançar valores impraticáveis (Anderson & Cassel, 1986).

Diversos estudos foram desenvolvidos com o objetivo de quantificar o número de amostras para estimar as propriedades do solo. Com poucas exceções, os resultados foram desanimadores (McBratney & Webster, 1983). Mulla & McBratney (2000) sugerem uma forma para estimar o número de pontos amostrais necessários para calcular, com a precisão desejada, o valor da sua média, desde que os valores sejam independentes e apresentem distribuição normal.

Para estimar parâmetros estatísticos sem conhecer a distribuição dos valores observados, uma das formas utilizadas é a técnica da reamostragem

aleatória, na qual cada observação de uma população tem a mesma probabilidade de ser incluída num conjunto reamostrado. A atual disponibilidade de recursos de informática permitiu que fossem desenvolvidas poderosas metodologias de reamostragem. A técnica “bootstrap” de estimação é uma delas e tem como pré-requisito que o número de valores observados é suficientemente grande para representar a população inteira.

A técnica “bootstrap” é um procedimento computacional intensivo de reamostragem, baseado na técnica da substituição, que possibilita estimar a distribuição amostral de estatísticas de interesse, tendo como base dados de uma amostra ou população. Utilizando esta técnica, é possível realizar comparações estatísticas em diversas áreas do conhecimento científico, tais como: economia, ecologia, genética, biologia, agronomia, etc, e suas subáreas (Efron & Tibshirani, 1993; Chung et al., 1996; Venkovsky et al., 1997; Amador et al., 2000; Jhun & Jeong, 2000). Aplicações da técnica “bootstrap” ao estudo da variabilidade de parâmetros hídricos do solo são encontradas nos trabalhos de Dane et al. (1986), Schaap & Leij (1998) e Hendrickx & Wierenga (1990). A grande vantagem do método “bootstrap” é que ele independe do tipo de distribuição de frequência das observações, tornando desnecessária a prévia normalização dos dados, como acontece com a maioria dos métodos convencionais.

O conhecimento da função $K(\theta)$ é essencial para a solução de problemas relacionados com o movimento de água no solo nas áreas das Ciências

Agrárias e Ambientais, e a principal dificuldade para o seu correto uso nos processos de movimento de água no campo é seu elevado grau de variabilidade (Warrick & Nielsen, 1980; Vieira et al., 1981; Falleiros et al., 1998; Jong van Lier & Libardi, 1999). Esta elevada variabilidade requer um grande número de repetições para que se obtenham valores médios com pequena margem de erro e baixo coeficiente de variação.

Os objetivos deste estudo foram: verificar a relação entre o número de repetições em um experimento de perfil instantâneo e o coeficiente de variação para os parâmetros da função $K(\theta)$, utilizando o método "bootstrap", conforme descrito por Efron & Tibsthirani (1993); comparar o método tradicional e o "bootstrap" para determinação do número de amostras necessárias para estimar a média dos fatores da função condutividade hidráulica e teor de água no solo.

MATERIAL E MÉTODOS

Aspectos teóricos

No método do perfil instantâneo, busca-se uma solução para a equação de Richards na direção vertical:

$$\frac{d\theta}{dt} = \frac{d}{dz} \left[K(\theta) \frac{d\phi_t}{dz} \right] \quad (1)$$

sujeita às seguintes condições de contorno:

$$L = 0; t \geq 0; q = 0$$

$$L = L; t > 0; q = q_L$$

Nesta equação, θ é a umidade volumétrica ($m^3 m^{-3}$), t é o tempo de redistribuição (h), z é a coordenada vertical de posição (m), L é a profundidade no solo (m), $K(\theta)$ é a função condutividade hidráulica ($m^2 h^{-1} kPa^{-1}$), ϕ_t é o potencial total da água no solo (kPa) e q é a densidade de fluxo da água no solo ($m h^{-1}$).

Integrando a equação de Richards (1) com relação à profundidade L , entre os limites $L = 0$ e $L = L$ obtém-se:

$$\int_0^L \frac{d\theta}{dt} dL = \int_0^L \frac{d}{dz} \left[K(\theta) \frac{d\phi_t}{dz} \right] dL \quad (2)$$

lembrando que $dL = -dz$ e $q = 0$ para $L = 0$:

$$-\int_0^L \frac{d\theta}{dt} dL = \left[K(\theta) \frac{d\phi_t}{dz} \right]_L = q_L \quad (3)$$

ou, rearranjando:

$$K(\theta)|_L = \frac{-\int_0^L \frac{d\theta}{dt} dL}{\frac{d\phi_t}{dz}|_L} = \frac{-\frac{dh_L}{dt}}{\frac{d\phi_t}{dz}|_L} \quad (4)$$

sendo h_L (m) a armazenagem de água no solo entre a superfície ($L = 0$) e a profundidade do solo L . Por meio de medidas simultâneas do teor de água e do potencial total ao longo do perfil de solo (de $L = 0$ a $L = L$) durante a redistribuição da água, determina-se a variação de armazenagem com o tempo t e o gradiente de ϕ_t para cada profundidade de L que aparece na equação (4) e, então, o valor de $K(\theta)$ para aquele teor de água, na profundidade escolhida.

Normalmente, a função $K(\theta)$ é uma relação exponencial entre K e θ (Hillel et al., 1972; Libardi et al., 1980), de maneira que ela pode ser expressa pela equação:

$$K = K^* e^{\beta\theta} \quad (5)$$

sendo β uma constante adimensional e K^* a condutividade hidráulica para $\theta = 0$. O valor de K^* não tem um significado físico, porém é importante para entender, juntamente com o β , a variabilidade da função $K(\theta)$. A equação (5) pode ser escrita na forma logarítmica:

$$\ln K = \beta\theta + \ln K^* \quad (6)$$

permitindo a obtenção dos parâmetros β e $\ln K^*$ por regressão linear. Observe que β é o coeficiente angular da relação $\ln K$ versus θ e $\ln K^*$, o linear.

Desenvolvimento experimental

Realizou-se um experimento no campus da Escola Superior de Agricultura "Luiz de Queiroz" - ESALQ/USP, localizado no município de Piracicaba (SP), nas coordenadas geográficas 22° 42' 30" de latitude Sul, 47° 38' 00" de longitude Oeste e altitude 580 m. Utilizou-se um Latossolo Vermelho-Amarelo distrófico textura média (Typic Hapludox), que se apresenta moderadamente a bem drenado e o uso atual é com pastagem de capim-elefante.

Na parcela experimental, com comprimento de 70 m e largura de 20 m, foram selecionados, ao longo do comprimento e na linha central, 45 pontos de observação, distanciados de 1 m entre si (Figura 1). Em cada um deles, foram instalados tensiômetros (cápsulas porosas nas profundidades de 0,30; 0,45 e 0,60 m) e, no ponto central de cada conjunto de tensiômetros, foi instalada uma sonda de TDR, para medir o teor de água nas mesmas três profundidades. Os manômetros de mercúrio e os cabos de leitura do TDR ficaram localizados fora da área útil experimental, a uma distância de aproximadamente 10 m dos pontos de observação.

As sondas de TDR utilizadas são especificadas como "Type" 4 (PRB-H), model MP 917 da empresa E.S.I. Environmental Sensors Inc., com cinco segmentos de 0,15 m para medida do teor de água.

Para a aplicação do método do perfil instantâneo, o solo da parcela experimental foi umedecido, aplicando-se água por meio de irrigação por aspersão, continuamente, durante cinco dias, até os

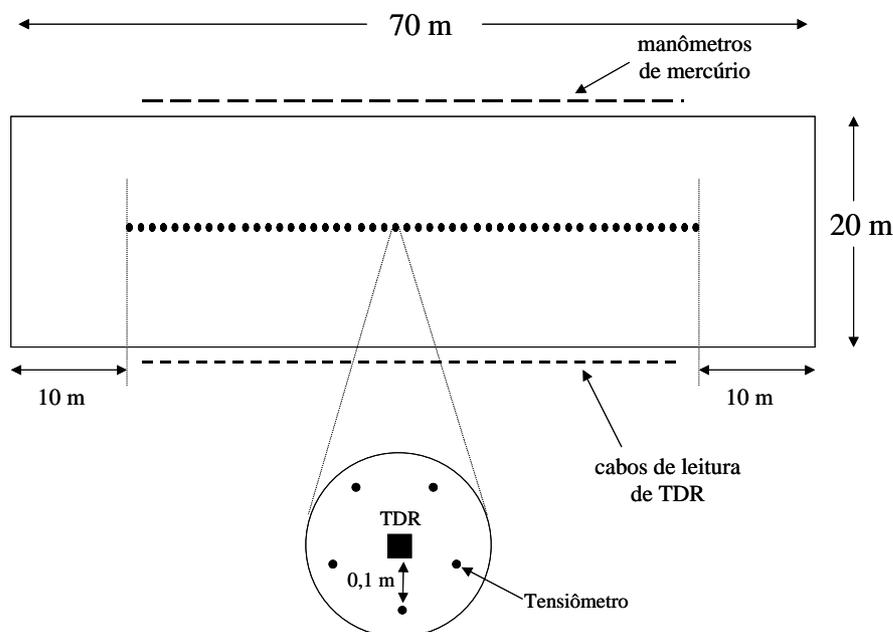


Figura 1. Parcela experimental com a localização dos 50 pontos de observação. O detalhe mostra a configuração de um ponto de observação com TDR e tensiômetros.

tensiômetros indicarem leitura constante. Em seguida, a superfície do solo foi coberta com uma lona plástica, satisfazendo, assim, a primeira condição de contorno exigida pelo método. A partir desse momento e durante o processo de redistribuição, o teor de água (θ) e o potencial mátrico da água (ϕ_m) no solo foram monitorados pelas leituras diárias das sondas do TDR e dos tensiômetros.

O período considerado para as análises e avaliações deste estudo ficou compreendido entre os tempos 78 e 604 horas de redistribuição da água, após a cobertura da superfície do solo com a lona plástica ($t = 0$), correspondentes à parte dos meses de junho e julho de 2000. Cinco pontos de observação foram eliminados em decorrência de defeitos nos TDRs e, ou, tensiômetros. As leituras do TDR foram calibradas em relação à umidade do solo mediante algumas observações simultâneas da leitura e da umidade volumétrica, determinada em amostras de volume conhecido retiradas com amostrador de Uhland.

Análise estatística dos resultados

O processamento dos dados experimentais, para determinação e cálculo da função $K(\theta) = K^* \exp(\beta\theta)$, foi feito, utilizando-se o programa de computador descrito por Jong van Lier & Libardi (1999). Como resultado, foram gerados 45 pares de valores de $\ln K^*$ e β para cada profundidade (sendo um par para cada ponto de observação), os quais foram submetidos à análise exploratória com vistas em avaliar o padrão de distribuição dos dados. Para tanto, utilizou-se o programa Statistica for Windows (Stat Soft, 1993).

As seguintes medidas estatísticas foram feitas: média, mediana, moda, desvio-padrão, variância, coeficiente de variação, valor mínimo e máximo, amplitude total, primeiro quartil, terceiro quartil, amplitude interquartil, assimetria e curtose, além da identificação de "outliers", segundo as indicações de Libardi et al. (1996). A distribuição dos dados quanto à sua normalidade foi feita com base nos coeficientes de assimetria e curtose, teste de Kolmogorov-Smirnov, análise visual da reta de Henry e "box-plot" e nas tabelas apresentadas por Jones (1969).

Santos & Vasconcelos (1987) propuseram a equação

$$n = \left(\frac{t_\alpha \cdot CV}{D} \right)^2 \quad (7)$$

para calcular o número mínimo de amostras necessárias para estimar o valor médio representativo de uma população, sendo n o número mínimo de amostras para estimar o valor médio representativo de uma variável, t_α o valor do teste t de Student para determinado nível de probabilidade, CV o coeficiente de variação e D a diferença desejada em torno da média. Com base nessa equação, estimou-se o número mínimo de amostras necessárias para obter o valor médio representativo de β e $\ln K^*$, 5 %, para os valores de coeficiente de variação de 5, 10, 20 e 30 % em torno da média.

Considerando, ainda, os dados do experimento, foi feita uma simulação *bootstrap* (Efron & Tibshirani, 1993), para medir o tamanho de amostra com base na média e variância, avaliando as estimativas de β e de $\ln K^*$. A técnica *bootstrap* é um método computacionalmente intensivo que se

baseia na reamostragem com reposição dos dados um grande número de vezes, sendo estimados, para cada reamostragem, os parâmetros de interesse. Com isso, torna-se possível obter uma distribuição empírica para estas estimativas *bootstrap*, bem como novas estimativas dos parâmetros de interesse, como a média, variância, intervalos de confiança, dentre outros. Essa técnica é chamada *bootstrap* não-paramétrico, uma vez que as estimativas *bootstrap* são baseadas apenas nos dados, não sendo esperada distribuição de probabilidade para esses dados.

No presente trabalho, foi desenvolvida uma rotina em S-Plus, v.3.3 (Statistical Science, 1993) e, com o uso do comando *sample*, foram geradas 10.000 replicações *bootstrap* para os dados experimentais com tamanhos de amostras n entre 5 e 45. Para cada subamostra, foram calculadas a média e a variância. Desta forma, 10.000 valores médios foram obtidos para cada n , permitindo estimar o coeficiente de variação das médias de acordo com o número de observações.

RESULTADOS E DISCUSSÃO

Estatística descritiva e análise exploratória dos dados

O quadro 1 apresenta os parâmetros estatísticos para todos os valores de β e $\ln K^*$ nas três profundidades consideradas neste estudo. Para melhor compreender e avaliar estes resultados, deve-se considerar primeiramente a influência de possíveis "outliers" sobre os parâmetros estatísticos que caracterizam os valores de β e $\ln K^*$ (Hamlett et al., 1986; Libardi et al., 1996). Identificou-se a presença de valores extremos nas três profundidades para ambos os coeficientes (Quadro 2). Os valores extremos influíram na média, na mediana, no desvio-padrão, no coeficiente de variação e nos índices de assimetria e curtose, resultando em uma distribuição não-normal para o conjunto de dados originais. A

eliminação dos valores extremos provocou mudanças importantes. Primeiramente, os valores da média e mediana tornaram-se mais próximos, principalmente nas profundidades de 0,30 e 0,45 m, alterando a forma da distribuição que passou a ser normal (Quadro 3). Admitindo que a condutividade hidráulica do solo apresente distribuição lognormal e que a sua transformação para $\ln K$ resulte em um conjunto de dados com distribuição normal (Nielsen et al., 1973; Libardi et al., 1980), esperava-se que a distribuição do conjunto de dados originais de $\ln K^*$ fosse aproximadamente normal, o que não ocorreu (Quadro 1). Tal fato deveu-se à existência dos "outliers" (Quadro 2), cuja presença no conjunto de dados foi responsável por alterar a distribuição dos valores da condutividade hidráulica, mesmo depois da aplicação da transformação adequada, evidenciando, mais uma vez, a importância da análise exploratória, como procedimento-padrão, na avaliação estatística de um conjunto de dados. Neste caso, a presença de valores extremos também alterou a distribuição dos dados, afastando a média da mediana e tornando-os não-normais (Quadro 3). A figura 2 mostra um exemplo para ilustrar o efeito da presença dos "outliers" na distribuição dos dados de β e de $\ln K$, na profundidade de 0,30 m. Nos outros casos, o comportamento foi semelhante.

Comprovada a hipótese da normalidade e assumindo a independência dos dados, calculou-se, com base na equação 7, o número de amostras necessárias para estimar o valor de β e $\ln K^*$ (Quadro 4). Pelo método tradicional, a profundidade de 0,45 m foi a que apresentou o menor número necessário de medidas para estimativa de β e $\ln K^*$. No caso de uma variação de 10 % em torno da média, por exemplo, seriam necessárias 12 medidas de β e 6 de $\ln K^*$. Nas outras profundidades, os valores são maiores. Analisando a equação 7, verifica-se que uma maior exigência, isto é, um menor D , resulta num número maior de observações, pois cada vez que D dobra, n fica quatro vezes menor.

Quadro 1. Resumo estatístico dos dados originais para valores de β e $\ln K^*$ da relação $\ln K = \beta\theta + \ln K^*$, em três profundidades, segundo uma transeção de 45 pontos espaçados de um metro (para K em $m^2 h^{-1} kPa^{-1}$)

Profundidade	Média	Mediana	Desvio-padrão	C.V.	Mínimo	Máximo	Assimetria	Curtose
m				%				
				β				
0,30	51,40	47,71	22,44	43,65	14,11	162,64	2,82	13,33
0,45	53,59	44,50	32,96	61,50	20,70	223,12	3,68	16,40
0,60	61,90	44,19	73,56	118,84	6,24	496,81	5,00	28,84
				$\ln K^*$				
0,30	-10,20	-9,49	5,59	54,80	-34,93	3,13	-1,77	8,38
0,45	-10,34	-8,52	7,46	72,15	-48,24	2,74	-3,36	15,42
0,60	-11,69	-8,12	16,46	140,80	-109,21	5,24	-5,00	29,11

Quadro 2. Limite inferior e superior para identificação de valores candidatos a “outliers” para os coeficientes (angular (β) e linear ($\ln K^*$)) da relação $\ln K = \beta\theta + \ln K^*$, em três profundidades, segundo uma transeção de 45 pontos espaçados de um metro (para K em $m^2 h^{-1} kPa^{-1}$)

Profundidade	1º Quartil	3º Quartil	Limite inferior	Limite superior	Número de “outliers”
m					
β					
0,30	39,89	57,81	13,01	84,70	(2)
0,45	37,76	52,79	15,21	75,34	(6)
0,60	34,26	57,68	-0,87	92,81	(5)
$\ln K^*$					
0,30	-12,37	-7,50	-19,68	-0,20	(3)
0,45	-10,84	-7,23	-16,25	-1,81	(6)
0,60	-11,51	-6,08	-19,65	2,06	(6)

Quadro 3. Resumo estatístico dos dados após a retirada dos “outliers” para valores de β e $\ln K^*$ da relação $\ln K = \beta\theta + \ln K^*$, em três profundidades, segundo uma transeção de 45 pontos espaçados de um metro (para K em $m^2 h^{-1} kPa^{-1}$)

Profundidade	Média	Mediana	Desvio-padrão	C.V.	Mínimo	Máximo	Assimetria	Curtose
				%				
m								
β								
0,30	47,84	47,58	13,33	27,87	14,11	72,80	-0,31	0,04
0,45	43,60	43,19	9,11	20,90	20,70	68,66	0,34	1,06
0,60	44,02	41,47	18,25	41,46	6,24	92,29	0,51	0,55
$\ln K^*$								
0,30	-9,86	-9,45	3,35	33,98	-16,53	-1,21	0,17	-0,10
0,45	-8,60	-8,36	2,28	26,51	-15,70	-3,27	-0,58	1,51
0,60	-8,12	-7,70	3,84	47,29	-17,48	-0,08	-0,38	0,60

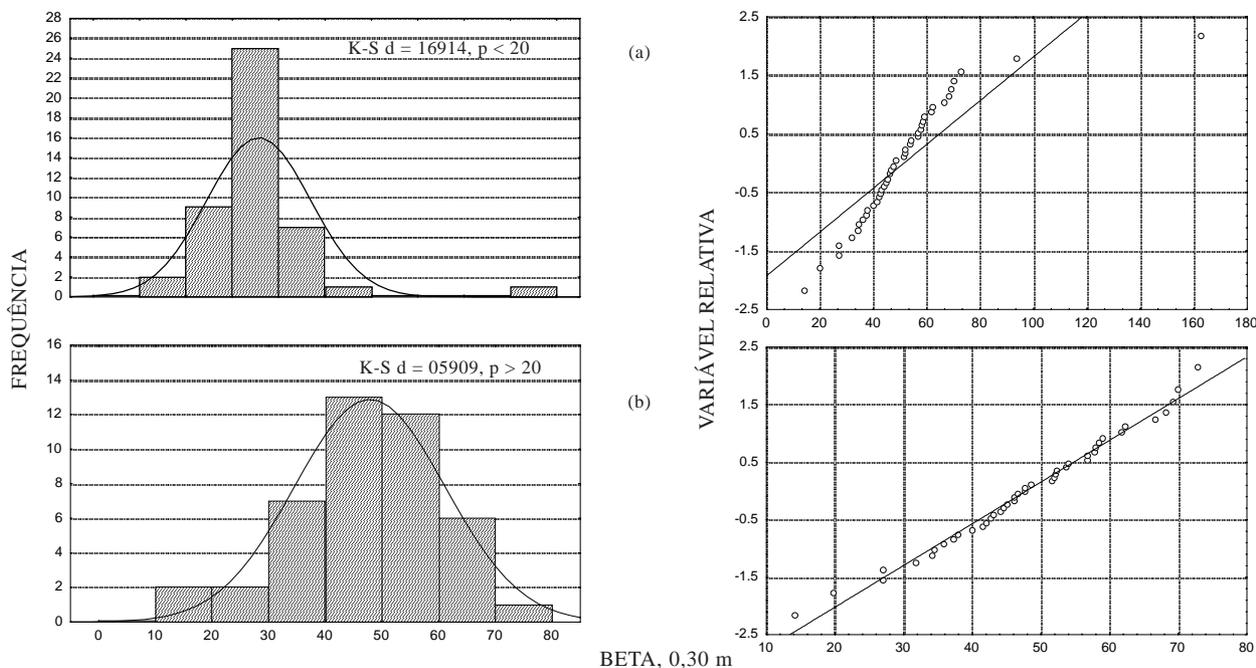


Figura 2. Histogramas de frequência, teste de Kolmogorov-Smirnov (K-S) e retas de probabilidade de Henry, dos valores originais, para o coeficiente angular (β) da relação $\ln K = \beta\theta + \ln K^*$, segundo uma transeção de 45 pontos espaçados de um metro, antes (a) e depois (b) da identificação e retirada dos valores “outliers”.

Quadro 4. Número de amostras necessárias para estimar o valor médio para o β e $\ln K^*$ da relação $\ln K = bq + \ln K^*$, em três profundidades, para quatro níveis de variação, considerando a de 10 % e utilizando o método convencional e o "bootstrap"

Profundidade	Convencional				"Bootstrap"			
	5	10	20	30	5	10	20	30
m	%							
	β							
0,30	88	22	5	1	> 45	> 45	12	5
0,45	49	12	3	2	> 45	> 45	23	9
0,60	194	49	12	5	> 45	> 45	> 45	34
	$\ln K^*$							
0,30	46	12	3	1	> 45	24	7	5
0,45	24	6	2	1	> 45	45	11	6
0,60	74	18	5	2	> 45	> 45	45	16

Análise pelo método "bootstrap" para β e $\ln K^*$

Como a análise pelo método "bootstrap" independe do tipo de distribuição dos dados, a normalização dos dados, ou a exclusão de "outliers", torna-se desnecessária. Os resultados da aplicação da técnica "bootstrap" para cálculo do número mínimo de amostras necessárias para estimar β e $\ln K^*$ encontram-se nas figuras 3 e 4, respectivamente. Os gráficos relacionam as frações (F) das 10.000 repetições da técnica "bootstrap" com variabilidade de $\pm 5, 10, 20, 30$ e 50% em torno da média. Em ambos os casos, β e $\ln K^*$, os resultados mostram que, para quase todos os níveis de variação em torno da média, ocorre um aumento das frações F, à medida que se eleva o número de amostras, pela técnica "bootstrap", até um patamar em que esse incremento torna-se tão pequeno que pode ser considerado estável.

Considerando um erro de $\pm 10\%$, que é aceitável para amostragem em agricultura (Hendrickx & Wierenga, 1990), as figuras 3 e 4 indicam que o número de observações necessárias para estimar o valor médio representativo de β e $\ln K^*$ pode alcançar valores muito mais elevados que 45 amostras, como nos casos de β , nas profundidades de 0,45 e 0,60 m, e $\ln K^*$, na profundidade de 0,60 m, o que tornaria impraticável a utilização desse nível de precisão (90 %) em amostragem da condutividade hidráulica. Assim, para o β , na profundidade de 0,45 m, o nível de erro aceitável para obtenção de um número factível de amostras seria 20 %, enquanto, na profundidade de 0,45, ter-se-ia de admitir um erro de 30 % para alcançar o mesmo objetivo. Quanto ao $\ln K^*$, o erro aceitável deve ser de, no mínimo, 20 %.

No quadro 4, onde se comparam esses valores com os obtidos pelo método tradicional, verifica-se que os valores obtidos pelo técnica "bootstrap" são bem maiores. Esse fato deve-se, provavelmente, à não-exclusão de "outliers", aumentando a variância do conjunto de dados.

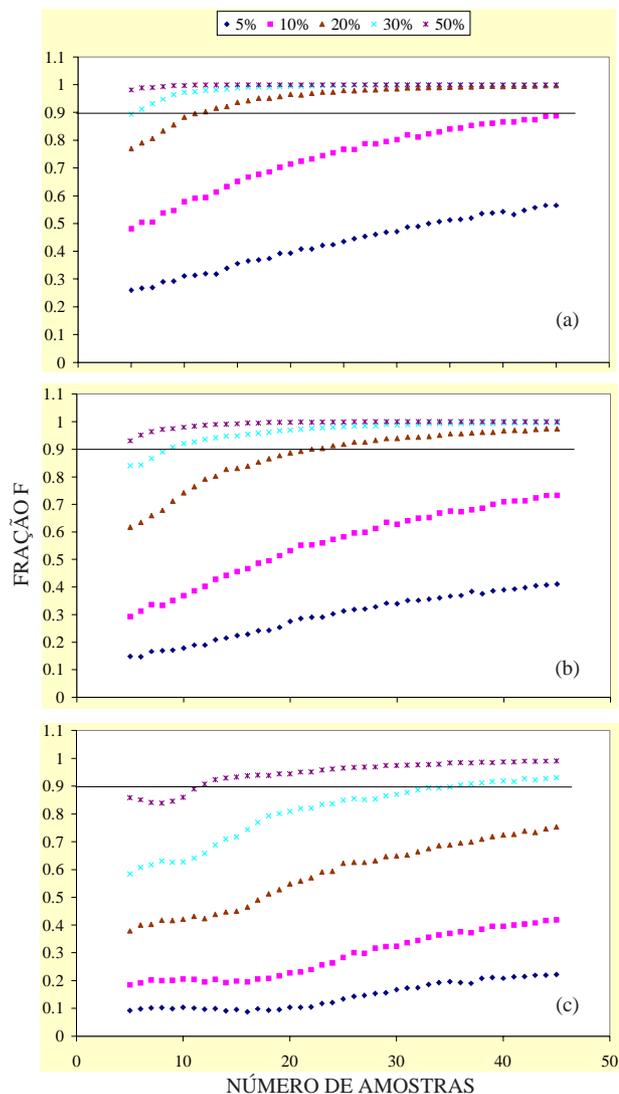


Figura 3. Fração da amostra "bootstrap" para β , considerando 5, 10, 20, 30 e 50 % de variação em torno da média e do tamanho da amostra, nas profundidades de 0,30 (a), 0,45 (b) e 0,60 m (c).

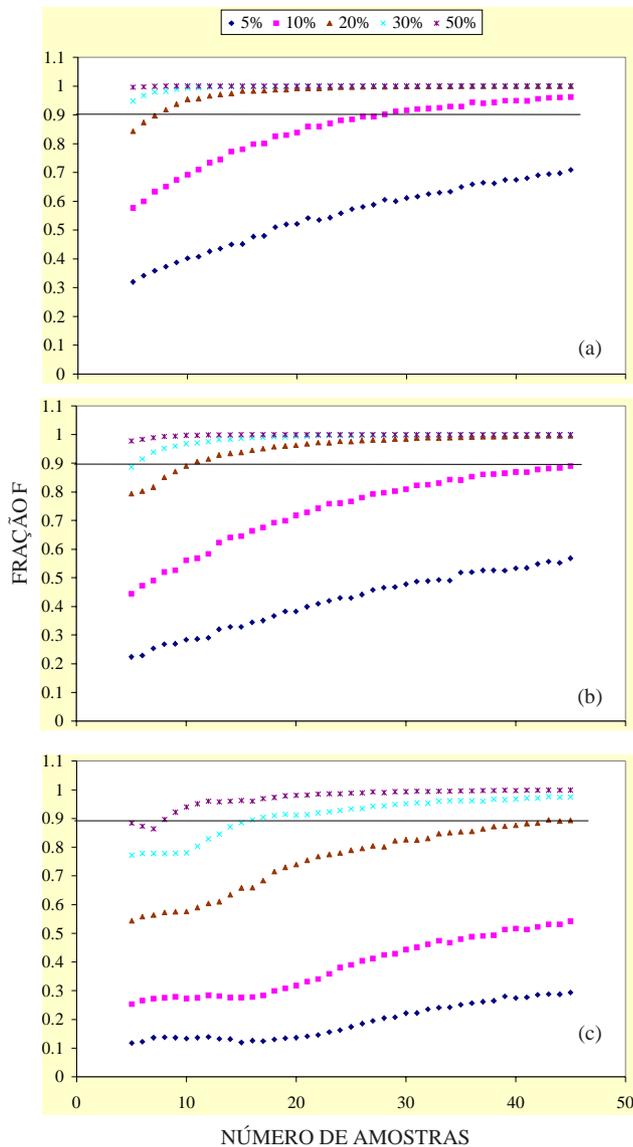


Figura 4. Fração da amostra “bootstrap” para $\ln K_{\theta=0}$, considerando 5, 10, 20, 30 e 50 % de variação em torno da média e do tamanho da amostra, nas profundidades de 0,30 (a), 0,45 (b) e 0,60 m (c).

Para β , o número de amostras aumenta com a profundidade. Para estimar a média, com 10 % de variação permitida, é necessário um número de amostras superior a 45 nas três profundidades. No mesmo nível de erro, 10 %, e nas mesmas profundidades, para o $\ln K^*$ o número de amostras também aumenta em profundidade e seriam necessárias 24, 45 e > 45 amostras. Considerando as características experimentais do método do perfil instantâneo, pode-se afirmar que com esses números torna-se difícil a condução do método em condições de campo, em que os valores de n são elevados e impraticáveis tanto para β quanto para $\ln K^*$. Isto também indica que os resultados de experimentos desta natureza, com poucas ou apenas uma

repetição, devem ser considerados com bastante ressalva e utilizados com muito critério, na medida em que são comprovadamente pontuais, confirmando as conclusões de Jong van Lier & Libardi (1999) também para as camadas mais internas do solo.

Tendo em vista o elevado grau de variabilidade da condutividade hidráulica não saturada (Warrick & Nielsen, 1980), variações em torno de 30 % podem ser admitidas como perfeitamente aceitáveis, o que permite indicar o número mínimo de 34 amostras para estimativa dos parâmetros da função $K(\theta)$, tendo em vista ser este o maior valor encontrado para β e $\ln K^*$ (profundidade de 0,60 m) considerando coeficiente de variação de 30 %, nas três profundidades estudadas neste experimento (Quadro 4).

Outra comparação muito oportuna deve ser feita entre os valores de n calculados pelo método tradicional, equação 7, e pela técnica “bootstrap” (Quadro 4). Nos dois casos, considerando variação igual ou superior a 10 %, o número de amostras estimadas pelo “bootstrap” foi sempre mais elevado. Não obstante, os dois métodos também mostraram semelhanças quanto à tendência de alguns resultados. Em ambos os casos, o número mínimo de amostras necessárias para estimar a média foi maior para o β . Por outro lado, a grande diferença entre os valores estimados mostra que a aplicação desses métodos de estimativa para médias de um conjunto de amostras relativas a propriedades, como a condutividade hidráulica, deve ser feita com certa prudência.

CONCLUSÕES

1. A comparação entre o método “bootstrap” e o convencional, que utiliza o valor t de Student para calcular o número de amostras (n) de β e $\ln K^*$, mostrou que as estimativas dos valores de n podem ser bastante diferentes. Como a técnica “bootstrap” independe do tipo de distribuição dos dados, é desnecessária a normalização dos dados, ou a exclusão de “outliers”, aumentando a credibilidade de seus resultados, razão por que é a mais indicada.

2. Ambos os métodos de avaliação resultaram em números elevados de repetições, reafirmando que determinações da função $K(\theta)$ pelo método do perfil instantâneo com poucas repetições não podem ser extrapoladas para áreas maiores. No caso do solo estudado, o número mínimo indicado pela técnica “bootstrap”, considerando três profundidades (0,30; 0,45 e 0,60), foi de 34 repetições.

AGRADECIMENTOS

Os autores agradecem à FAPESP, pelo auxílio concedido para a execução deste projeto de pesquisa (processo número 99/02130-0).

LITERATURA CITADA

- AMADOR, J.A.; WANG, Y.; SAVIN, M.C. & GÖRRES, J.H. Fine-scale spatial variability of physical and biological soil properties in Kingston, Rhode Island. *Geoderma*, 98:83-94, 2000.
- ANDERSON, S.H. & CASSEL, D.K. Statistical and autoregressive analysis of physical properties of Portsmouth sandy loam. *Soil Sci. Soc. Am. J.*, 50:283-287, 1986.
- CHUNG, H.; LEE, K. & KOO, J. A note on bootstrap model selection criterion. *Stat. Prob. Letters*, 26:35-41, 1996.
- DANE, J.H.; REED, R.B. & HOPMANS, J.W. Estimating soil parameters and sample size by bootstrapping. *Soil Sci. Soc. Am. J.*, 50:283-287, 1986.
- EFRON, B. & TIBSHIRANI, R.J. An introduction to the bootstrap. London, Chapman and Hall, 1993. 436p.
- FALLEIROS, M.C.; PORTEZAN, O.; OLIVEIRA, J.C.M.; BACCHI, O.O.S. & REICHARDT, K. Spatial and temporal variability of soil hydraulic conductivity in relation to soil water distribution, using an exponential model. *Soil Technol.*, 45:279-285, 1998.
- HAMLET, J.M.; HORTON, R. & CRESSIE, N.A.C. Resistant and exploratory techniques for use in semivariogram analysis. *Soil Sci. Soc. Am. J.*, 50:868-875, 1986.
- HILLEL, D.; KRENTOS, V.D. & STYLIANOU, Y. Procedure and test of an internal drainage method for measuring soil hydraulic characteristics in situ. *Soil Sci.*, 114:395-400, 1972.
- HENDRICKX, J.M.H. & WIERENGA, P.J. Variability of soil water tension in a trickle irrigated Chile pepper field. *Irrig. Sci.*, 11:23-30, 1990.
- JHUN, M. & JEONG, H. Applications of bootstrap methods for categorical data analysis. *Comp. Stat. Data Anal.*, 35:83-91, 2000.
- JONES, T.A. Skewness and kurtosis as criteria of normality in observed frequency distributions. *J. Sedim. Petr.*, 39:1322-1327, 1969.
- JONG VAN LIER, Q. & LIBARDI, P.L. Variabilidade dos parâmetros da relação entre condutividade hidráulica e umidade do solo determinada pelo método do perfil instantâneo. *R. Bras. Ci. Solo*, 23:1005-1014, 1999.
- LIBARDI, P.L.; REICHARDT, K.; NIELSEN, D.R. & BIGGAR, J.W. Simplified field methods for estimating the unsaturated hydraulic conductivity. *Soil Sci. Soc. Am. J.*, 44:3-6, 1980.
- LIBARDI, P.L.; MANFRON, P.A.; MORAES, S.O. & TUON, R.L. Variabilidade da umidade gravimétrica de um solo hidromórfico. *R. Bras. Ci. Solo*, 20:1-12, 1996.
- McBRATNEY, A.B. & WEBSTER R. How many observations are needed for regional estimation of soil properties? *Soil Sci.*, 135:177-183, 1983.
- MULLA, D.J. & McBRATNEY, A.B. Soil spatial variability. In: SUMNER, M.E., ed. *Handbook of soil science*. New York, CRC Press, 2000. p.321-352.
- NIELSEN, D.R.; BIGGAR, J.W. & ERH, K.T. Spatial variability of field measured soil water properties. *Hilgardia*, 42:215-259, 1973.
- SANTOS, H.L. & VASCONCELOS, C.A. Determinação do número de amostras de solo para análise química em diferentes condições de manejo. *R. Bras. Ci. Solo*, 11:97-100, 1987.
- SCHAAP, M.G. & LEIJ, F.J. Using neural networks to predict soil water retention and soil conductivity. *Soil Till. Res.*, 47:37-42, 1998.
- STAT SOFT. *Statistica for Windows*, v.4.3. Tulsa, Stat Soft, 1993.
- STATISTICAL SCIENCE. *S-Plus for Windows: user's manual*. Seattle, 1993. 2v.
- VIEIRA, S.R.; NIELSEN, D.R. & BIGGAR, J.W. Spatial variability of field-measured infiltration rate. *Soil Sci. Soc. Am. J.*, 45:1040-1048, 1981.
- VENKOVSKY, R.; DIAS, C.T.S.; DEMÉTRIO, C.G.B.; LEANDRO, R.A. & PIEDADE, S.M.S. Reamostragem por "bootstrap" na estimação de parâmetros baseados em marcadores genéticos. In: ENCONTRO SOBRE TEMAS DE GENÉTICA E MELHORAMENTO, 14., Piracicaba, 1997. *Anais. Piracicaba, ESALQ*, 1997. p.59-72.
- WARRICK, A.W. & NIELSEN, D.R. Spatial variability of soil physical properties in the field. In: HILLEL, D., ed. *Applications of soil physics*. New York, Academic Press, 1980. p.319-344.

