COKRIGING FIELD MEASURED SOIL HYDRAULIC CONDUCTIVITY AND TEXTURE IN A BRAZILIAN SEMI-ARID WATERSHED

Montenegro, A.A[(1)]; Lundgren, W.[(2)]; Montenegro, S.M.G.L.[(3)]

[(1)] Federal Rural University of Pernambuco State, Rural Technology Department, Brazil, [(2)] Federal Rural University of Pernambuco State, Mathematics Department, Brazil, [(3)] Federal University of Pernambuco State, Civil Engineering Department, Brazil

Abstract

The spatial relationship between saturated hydraulic conductivity and soil texture has been analysed, in a semi-arid watershed in the Brazilian Northeast. Adopting a 30m spacing regular mesh, 49 hydraulic conductivity tests have been carried out using a Guelph permeameter. Trend components have not been detected on data. A transformation was applied to the hydraulic conductivity data to obtain Normality. Isotropic semivariograms for the transformed hydraulic conductivity exhibited weak spatial dependence for the adopted spacing, whereas for texture large ranges for the semivariograms have been observed. Although linear correlation between hydraulic conductivity and texture was weak, the spatial correlation is strong. Adopting the cokriging procedure for hydraulic conductivity and texture, consistent maps could be produced, using the ArcView Software. The use of the cross-semivariogram has allowed approximately 40% reduction in the number of hydraulic measurements required, without increasing the precision error of the estimates.

Introduction

When modelling infiltration in hillslope areas, knowledge of the hydraulic conductivity parameter variations in the soil zone is of critical importance for accurately predicting runoff and seepage events. Even in a detailed soil investigation in watersheds, many areas between sampling points are left unmeasured, and some degree of uncertainty remains concerning the actual rainfall- runoff process. This uncertainty can be quantified by treating the soil parameters as random functions (de Marsily, 1986). Given a set of measurements at specified field locations, geostatistical analysis consists of inferring (in a statistical sense) the parameters of the probability model governing the behaviour of each random function in space (Mackay et al., 1987). In addition, the geostatistical approach allows the generation of many realisations of this same random function, of which the actual measurements represent only one particular realisation.

Comprehensive literature review has been presented by Mackay et al. (1987) on the geostatistical characteristics of hydraulic conductivity in several unconsolidated formations, regarding particularly on correlation length, as related to the size of the study region. Ranges in alluvial soils have been reported from 100m to 800m.

In this study, a hillslope plot in semi-arid watershed in the Brazilian Northeast has been selected and its hydraulic properties structure analysed. In the same plot, Santos et al. (2003) have performed rainfall-runoff experiments in the selected plot using a rainfall simulator device. Values between 2cm/h and 8cm/h have been accessed for saturated hydraulic conductivity during the dry season, when bare soil scenarios prevail.

Mackay and Montenegro (1996), Montenegro (1997) and Montenegro et al. (2002) have investigated the geostatistical structure of hydraulic conductivity in the same watershed, particularly in the alluvial valley downstream from the hillslope plot adopted. Ranges of approximately 84 m for the isotropic semivariogram have been found, for the saturated hydraulic conductivity.

In order to investigate spatial variability of hydraulic conductivity, a Guelph permeameter (Soil Moisture, 1991) has been applied in this study, due to its versatility, allowing hydraulic conductivity assessment in about 2 hours time interval, on average. According to Ragab and Cooper (1993), the Guelph permeameter should be used where shallow water tables are not present, producing precise hydraulic conductivity estimates, and causing minimal disturbances onto soil structure. Ragab and Cooper (1993) have described an application of the Guelph permeameter to compare unsaturated hydraulic conductivity in grasslands and forest zones in a same watershed. When applying the permeameter in sandy soils, borehole wall collapsing might occur. Bagarello (1997) has inserted a wire screen to prevent collapsing, producing consistent results. This method has been adopted in this study, due to the sandy texture of the domain.

As pointed out by Vauclin et al. (1983), in several practical situations the spatial structure of a given variable may not been fully described from an univariate sampling procedure. The statistical precision may be improved by potentially considering the spatial correlation between this variable and another more frequently sampled one, as stated by the cokriging procedure.

The purpose of this paper is to apply the cokriging procedure to hydraulic conductivity evaluation and texture data from a semi-arid watershed, and to compare the results to classical kriging procedure.

Theory

Application of the geostatistical approach

Geostatistical theory is based on the assumption that the variability of a given regionalized variable **z(x)** is random in space. According to Journel and Huijbregts (1978), a random function **Z(x)** can be seen as a set of random variables $\mathbf{Z(x_i)}$ defined at each point $\mathbf{x_i}$ in the domain of interest **D**: $\mathbf{Z(x) = \{Z(x_i), \forall x_i \in D\}}$. The random variables $\mathbf{Z(x_i)}$ are spatially correlated, depending on the structure of variation of the random function **Z(x)**. At any point $\mathbf{x_i}$, the log-transformed value of the measured hydraulic conductivity, $\mathbf{z(x_i)}$, for example, can be interpreted as a single realization of the random variable $\mathbf{Z(x_i)}$. Assuming hypothesis of stationarity, as stated by Journel and Huijbregts (1978), the function $\boldsymbol{\gamma}$**(h)** is called the semi-variogram, and can be defined as:

$$\gamma(h)= 0.5 \{ E[(Z(x+h)-Z(x))^2] \} \quad (1)$$

where **E[.]** represents the mathematical expectation.

When $\boldsymbol{\gamma(h) \rightarrow c_0 > 0}$, **as** $\mathbf{h \rightarrow 0}$, then $\mathbf{c_0}$ is called the nugget effect, and incorporates measurement errors and small-scale variability not captured by the field sampling design (Cressie, 1993). In general, the semi-variogram can be interpreted as representing a nested structure of variation, and encompassing contributions of several scales of variability of a natural phenomenon. Each scale of observation integrates the variability of all the smaller scales (Journel and Huijbregts, 1978). The sill of the semi-variogram, $\boldsymbol{\gamma_2(\infty)}$, represents the limit of the $\boldsymbol{\gamma_2}$ **(h)** function as $\mathbf{h \rightarrow \infty}$, and the range λ of the semi-variogram indicates the separation scale at which the semivariogram reaches the sill.

Prior to the semi-variogram analysis from the measurement data, assessment of the univariate distribution for the probability model is required. According to Woodbury and Sudicky (1991), hydraulic conductivity values are commonly assumed to follow a lognormal distribution. In this study, lognormality tests were performed on the set of hydraulic conductivity measurements.

After assessing the probability density function for the random function **Z(x)**, the covariance structure of **Z(x)** can be investigated by invoking the intrinsic hypothesis and adopting an estimator function for the semi-variogram $\boldsymbol{\gamma}$**(h)**, for example the ‘classical’ experimental semi-variogram estimator $\boldsymbol{\gamma^*}$ **(h)**.

Suppose now that two variables Z_1 and Z_2, fulfilling stationarity requirements, have cross covariances different from zero, then the cross semivariogram can be estimated by:

$$\gamma_{12}^{*}(h_j)=\frac{1}{2n_j}\sum_{i,k=1}^{n}[z_1(x_i)-z_1(x_k)].[z_2(x_i)-z_2(x_k)] \quad (2)$$

$$z_{1,2}(x_{i,k})\in(h_j^- \le dist(z_{1,2}(x_i),z_{1,2}(x_k)<h_j^+)$$

where **n** is the total number of measurements, $\mathbf{n_j}$ is the number of pairs in the interval [$\mathbf{h_j^-}$ **;** $\mathbf{h_j^+}$], and **dist(z(x$_i$), z(x$_k$))** is the Euclidean distance in **D**.

In this study, the Weighted- Least- Square Method has been employed, as recommended by Cressie (1985), to fit theoretical semivariogram models to the experimental data. The most common functions used as semi-variogram models are the exponential, the spherical, the linear and the Gaussian (Journel and Huijbregts, 1978)

The performance of the theoretical semi-variogram model in characterizing the structure of a given data set should be verified before its use in mapping and prediction. The method of jack-knifing or cross-validation is commonly applied for such a validation (Vauclin et al., 1983), which consists of successively removing one data point z_i at a time and then estimating it from the remaining data, producing a value $\mathbf{z_i^*}$. The kriging technique is used in the estimation of $\mathbf{z_i^*}$. Thus, at each data location $\mathbf{x_i}$, a reduced error $\mathbf{R_i}$ can be calculated, which should be close to zero if no bias is present, and a reduced variance should be close to 1, suggesting that the variance of the errors $(\mathbf{z_i - z_i^*})$ is consistent with the kriging variance $(\sigma_{k\,i})^2$**.**

For the purpose of this study, the experimental semi-variogram for the log-transformed hydraulic conductivity was estimated based on the classical experimental semi-variogram estimator proposed by Matheron (1963). The class intevals were chosen such that at least 40 pairs of points were present in each class and the semi-variogram captured the small scale variation of the hydraulic conductivity field.

Kriging procedure description is well known in literature, representing a weighted moving average with an estimation of the form:

$$z^*(x_0)=\sum_{i=1}^{N}\lambda_i z(x_i) \quad (3)$$

where N is the number of measured values involved in the interpolation to x_0 value, and λ_i are weights. If $z(x_i)$ is a realization of the random function $Z(x_i)$ fufilling the intrinsic assumption, then equation (3) can be rewritten in terms of Z. It should be highlighted that one of the conditions for optimizing the choice of the λ_i is:

$$\sigma_K^2(x_0)=Var\{Z^*(x_0)-Z(x_0)\}=\min \quad (4)$$

For analyzing the cokriging variance, let z_2 be a more difficult variable to measure than z_1. Supposing that the two variables have the cross correlation expressed by $\gamma_{12}^*(h)$, Z_2^* can be estimated as:

$$Z_2^*(x_0) = \sum_{i=1}^{N1} \lambda_{1i} Z_1(x_{1i}) + \sum_{j=1}^{N2} \lambda_{2j} Z_2(x_{2j}) \quad (5)$$

The variance requirement is then:

$$\sigma_{cK}^2(x_0) = Var\{Z_2^*(x_0) - Z_2(x_0)\} = \min \quad (6)$$

In this study, mapping is performed for σ_K^2 and for σ_{cK}^2 aiming to assess the improvement in Z_1 estimates when cokriging is adopted. For more details, Vauclin et al. (1983) can be consulted.

Hydraulic conductivity tests

The Guelph permeameter has been employed adopting four hydraulic heads in boreholes (2.5; 5.0; 7.5 and 10.0 cm), allowing six pairs of heads to be assessed. Reynolds e Elrick (1986) have present the details for Gardner function parameters estimation from each pair of heads. The basic principle is that the steady state regime is quickly established in a borehole in a homogeneous isotropic domain, with a low hydraulic head.

The matric flux potential ϕ_m can be defined as:

$$\phi_m = \int_{\varphi_i}^{0} K(\varphi) d\varphi \quad (7)$$

$\varphi_i < \varphi < 0$, where, φ_i is the initial soil tension and φ is the final one.

Adopting the Gardner (1958) expression, where α is used as a soil porosity, the saturated hydraulic conductivity K_0 can be expressed as:

$$\frac{\phi_m}{K_0} = \alpha^{-1} \quad (8)$$

Results and discussion

A regular mesh with 30m spacing and 49 nodes has been adopted in the hillslope plot. In each node, texture has been evaluated and hydraulic conductivity assessed, with at least three replications. The data basic statistics is presented in Table 1. Figure 1 exhibits the Normal theoretical function and the experimental cumulative statistics for Log K. It has been verified that the sand content and the clay content are approximately Normally distributed, whereas hydraulic conductivity can be assumed as log-Normally distributed.

Table 1- Basic statistics for log K, sand content, and clay content.

Variable	n	max	min	average	sd	Average/sd
Log K(cm/h)	49	2.10	0.12	1.40	0.40	28.57
% sand	49	68.20	39.40	54.46	7.07	12.98
% clay	49	39.00	17.40	27.53	5.46	19.83

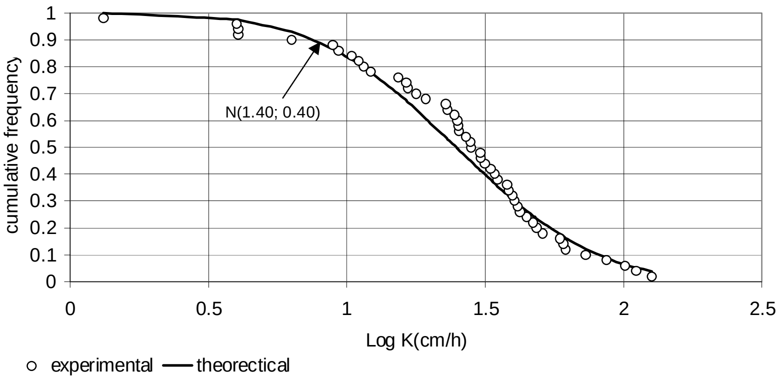

Figure 1- Cumulative frequency of Log (K).

The trend analysis on data has been performed by calculating the determination coefficient of the multiple regression of log K in both directions X and Y, as:

Log K= 1.07+0.0008X+0.002Y R^2= 0.13

As R^2 is far from unit, and also the linear coefficients, then it has been assumed that no trend is present in the data set.

Considering the hydraulic tests replication in each location in the field, the nugget effect in the semivariogram for Log K has been estimated in 0.04. Figure 2 exhibits the experimental and theoretical semivariances for texture data, and Figure 3 presents the cross-semivariograms log K – sand content and log K – clay content. Table 3 summarizes the optimal fitting of the semivariograms, and also presents the theoretical semivariogram after removing 19 sampling points from the mesh, aiming to evaluate the potential of cokriging procedure. Part of the removed points comes from an entire transect.

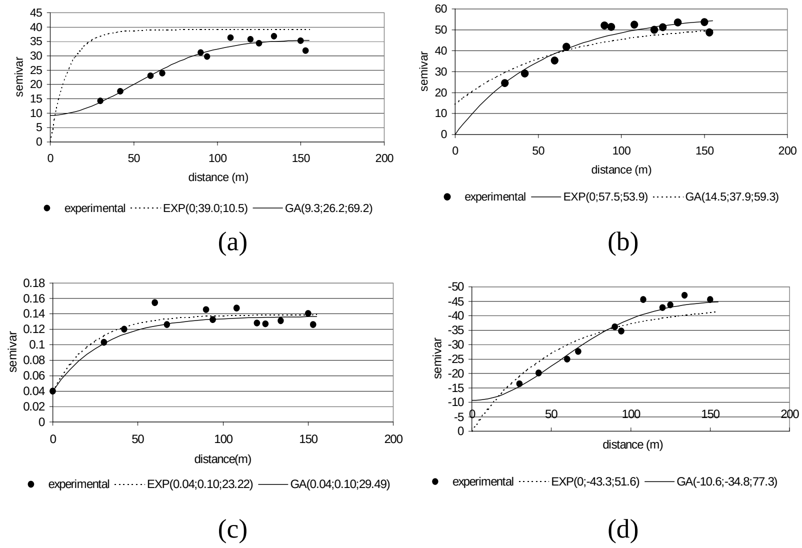

Figure 2- Semivariograms for sand content(a); clay content(b); log K (c) and the cross semivariogram clay-sand (d)

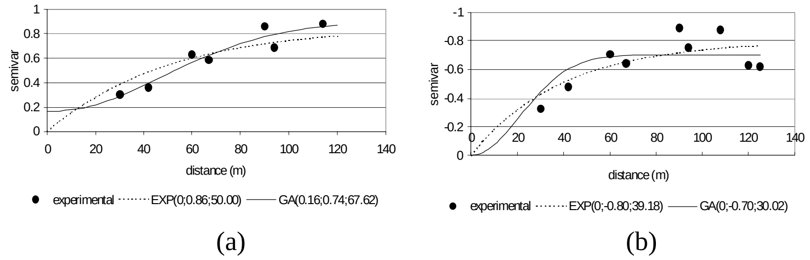

Figure 3- Cross semivariograms for logK-sand (a); and log K-clay (b).

Arcgis software (ESRI, 2001) has been employed to krige the 30 sampling points, and to cokrige the 30 hydraulic conductivity estimates with sand and clay contents. As the software limits the use of two distints semivariogram models while cokriging, adjustments had to be made to the clay content model.

Table 3- Summary of the optimal semivariograms fitting, for 49 and 30 nodes.

Variables	49 sampling points				30 sampling points			
	nugget	range	sill	model	nugget	range	sill	model
Clay-clay	0.00	161.70	57.50	E*				
Clay-sand	-10.61	133.89	-45.41	G**				
Sand-sand	9.30	119.86	35.50	G**				
log K – log K	0.04	51.07	0.14	G**	0.04	52.17	0.15	G**
log K - sand	0.16	117.12	0.90	G**	0.12	115.46	0.85	G**
log K - clay	0.00	52.00	-0.70	G**	0.00	50.34	-0.83	G**

*Exponential Model. (EXP).
** Gaussian Model. (GA).

Figure 4 presents the comparison of the standard error distribution and prediction errors for ordinary kriging and ordinary cokriging using 30 sampling points.

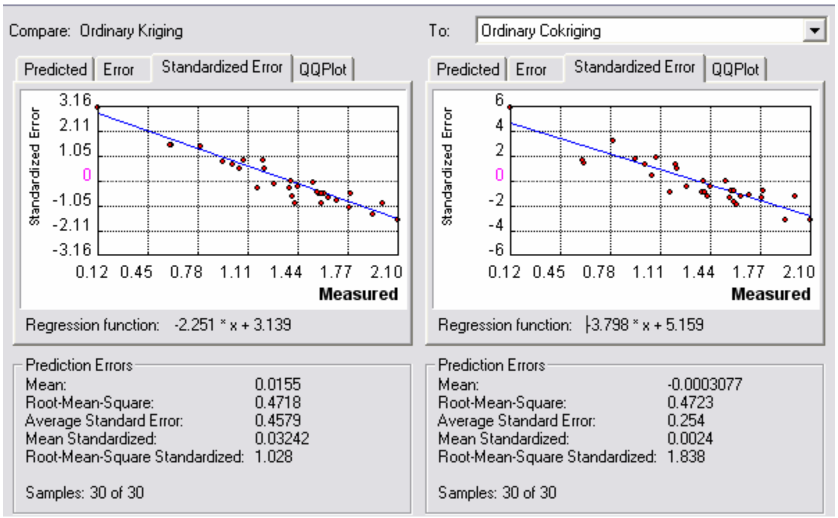

Figure 4. Standard error distribution and prediction errors for ordinary kriging.

The spatial distribution of prediction errors can be seen in Figure 5, where the circles are the 19 sampling points removed from the original mesh. The large ranges for sand and clay contents have contributed for reducing and smoothing the prediction errors.

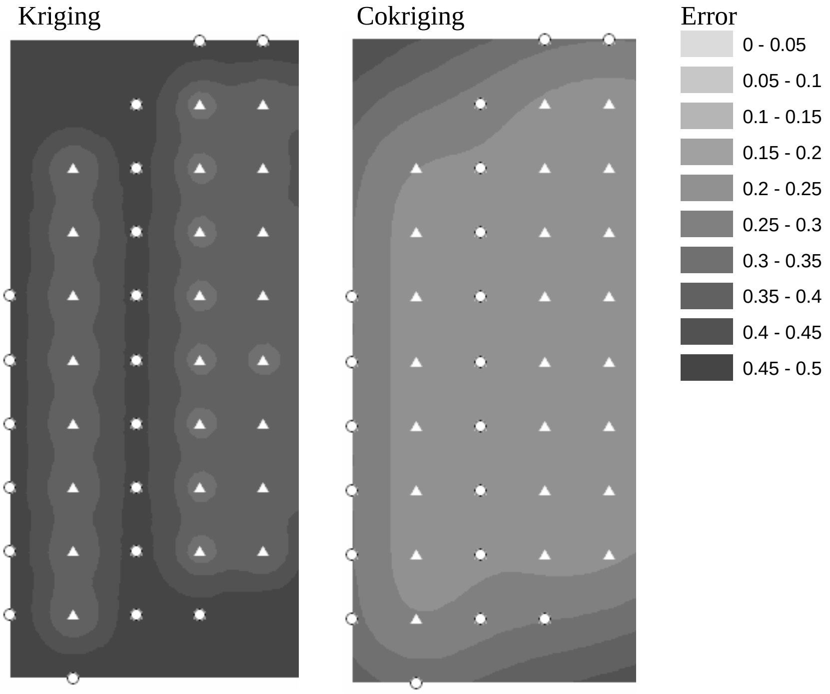

Figure 5. Prediction error map for kriging and cokriging estimation.

Figure 6 presents the kriging and cokriging standard deviations. It can be verified that the cokriging procedure have strongly reduced the standard deviation of the interpolation process.

The kriged scenario (z^*) of $Log_{10}K$(cm/h) is shown in Figure 7.

Kriging Cokriging

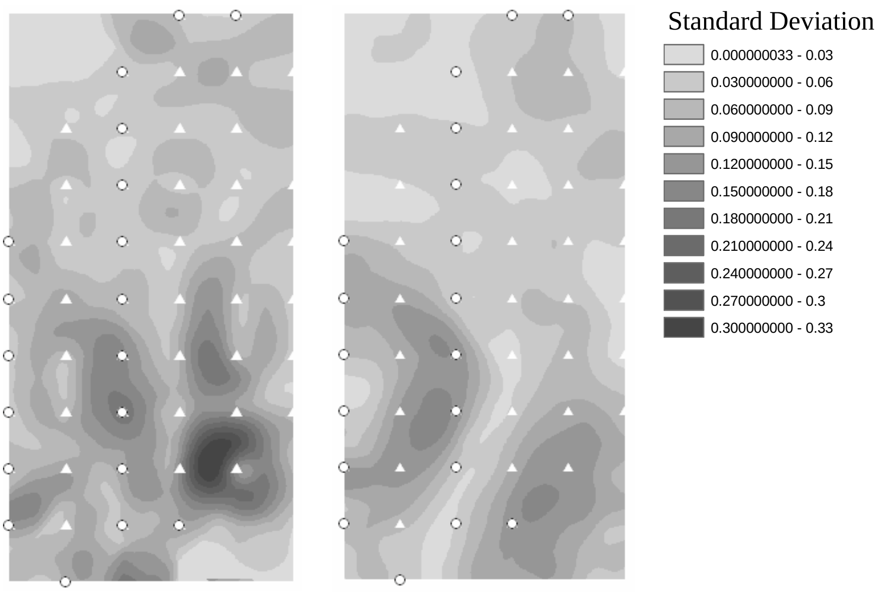

Figure 6. Standard deviation maps for kriging (a)and cokriging (b).

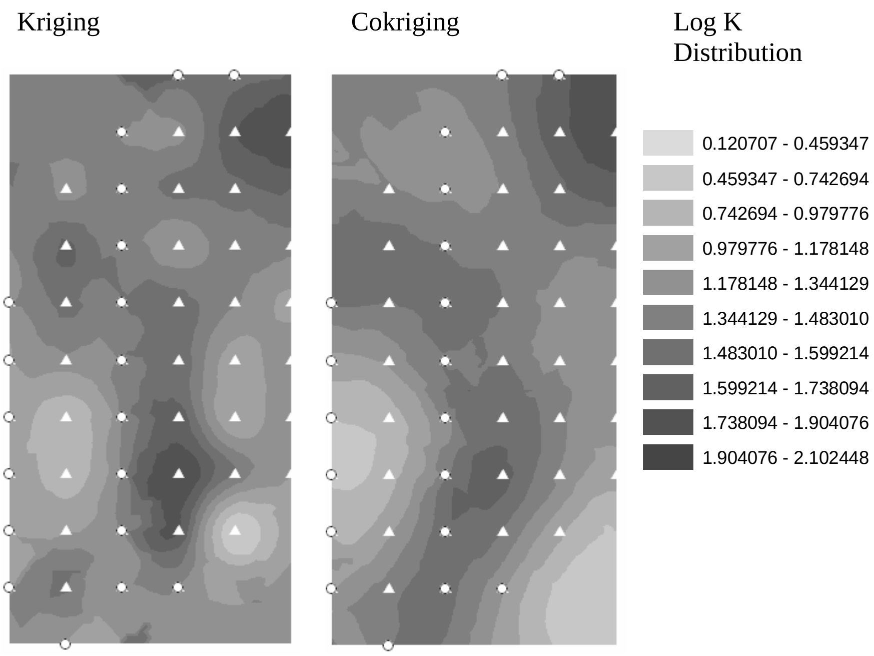

Figure 7. Spatial distribution for Log(K) using kriging and cokriging methods.

Conclusions

In this study, performed in a watershed hillslope in the brazilian semi-arid, it has been verified the relevance of cokriging procedure for mapping hydraulic conductivity, which is an expensive variable to assess in the field, based more frequently sampled variables, as sand and clay contents. For the studied domain, such soil fractions exhibit larger ranges than the hydraulic conductivity, and also present spatial correlation to hydraulic permeability. Texture were found to be normally distributed, while hydraulic conductivity distributed Log-Normally over the field. Kriging and cokriging have been used for spatial prediction of saturated hydraulic conductivity at a 30 points approximately regular mesh, with 30 nodes, obtained from a 49 nodes regular mesh by removing a transect and some random points. Cokriging procedure reduced the predition error of the estimates, and constitutes an important tool to provide, at unrecorded points, an unbiased estimation of Log K parameter with a minimum standard deviation.

Acknowledgments

This work has done thanks to financial support from FINEP and CNPq, brazilian federal Agencies for research. The authors also would like to acknowledge Eng. Thaísa Alcoforado de Almeida, for her suport in analysing and editing this paper.

References

Bagarello, V. Influence of well preparation on field saturated hydraulic conductivity measured with the Guelph permeameter. Geoderma, v. 80 p., 169 – 180, 1997.

Cressie, N. A. C. 1985. Fitting models by weighted least squares. *J. Math. Geology*, Vol. 17, No. 5: 563- 586.

Cressie, N. A. C. 1993. *Statistics for Spatial Data*. Revised edition. Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc. USA. 900p.

ESRI Inc. Arc Map 8.1. Copyright – US patent No. 5710835.2001.

Gardner, W. R. Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from water table. Soil Science, v. 85, p. 228 – 232, 1958.

Mackay, R., T. A. Cooper, J. D. Porter, P. E. O'Connell, and A. V. Metcalfe. 1987. *Geostatistical analysis of prevailing groundwater conditions and potential solute migration at Elstow, Bedfordshire*. Final Report to Associated Nuclear Services Ltd. England. 177p.

Mackay, R. and A. A. A. Montenegro. 1996. *Salinity control for sustainable small scale agriculture*, Final Report to the Overseas Development Administration (ODA), UK. 30p.

de Marsily, G. 1986. *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*. Academic Press, Inc. , Orlando, Florida. 440p.

Matheron, G. 1963. *Principles of Geostatistics*. Economic Geology, Vol. 58: 1246- 1266.
Journel, A. G. and Ch. J. Huijbregts. 1978. *Mining Geostatistics*. Academic Press. London. 600p.
Montenegro A. A., 1997: Stochastic hydrogeological modelling of aquifer salinization from small scale agriculture in northeast Brazil, PhD Thesis, University of Newcastle Upon Tyne, UK.
Montenegro, A. A. de A.; Mackay, R.; Montenegro, S. M. G. L. Coupled unsaturated saturated modelling of salinization risk in na alluvial irrigated área of North-East Brazil. *Acta Universitatis Carolinae Geologica*, v. 46, n. 2/3, p. 593- 594, 2002.
Ragab, R. ; Cooper, J. D. Variability of unsaturated zone water transport parameters: implications for hydrological modelling. 1. In situ measurements. Journal of Hydrology. p. 109 – 131, v. 148, 1993.
Reynolds, W. R. ; Elrick, A. A. Method for simultaneous in situ measurement in the valdose zone of field saturated hydraulic conductivity, sorptivity and conductivity – pressure – head relationship. Ground Water Monit., v. 6, p. 84 – 95, 1986.
Santos, T. E. M. dos; Oliveira, O. F. L.; Nascimento, J. J.; Montenegro, A. A. A.; Cantalice, J. R. B. Avaliação do Impacto da Exploração Agrícola nas Taxas de Erosão de uma Vertente em Agroecossistema Semi-árido. In: III Jornada de Ensino Pesquisa e Extensão, UFRPE, Recife- Brazil, 2003, 4p. CD- ROM.
Soil Moisture. Guelph Permeameter Operating Instructions Manual. Santa Bárbara: Soil Moisture equip. corp., 1991.
Vauclin M., Vieira S. R., Vachaud G., Nielsen D. R., 1983: The use of cokriging with limited field soil observations. *Soil Sci. Soc. Am. J.*, 47: 175-184.
Woodbury, A. D.; Sudicky, E. A. The geostatistical characteristics of the Borden Aquifer , Water Resources Reasearch, v. 77, n. 4, p. 533 - 546, 1991.