Estimating and Simulating Spatial Distribution of Transmissivity Using Geostatistics Methods

Yu-Pin Lin [a,b], Tung-Po Teng [b], Yih-Chi Tan [c]

[a]*Department of Landscape Architecture, Chinese Culture University, 55 Hwa-Ken Rd. Yangming Shan , Taipei, 111 Taiwan (yplin@staff.pccu.edu.tw)*

[b]*Graduate Institute of Geography, Chinese Culture University.*

[c]*Graduate Institute of Agricultural Engineering, National Taiwan University.*

Abstract: Groundwater modelling requires proper and accurate hydrogeological properties such as transmissivity and hydraulic conductivity. This study identified the spatial patterns and variations in transmissivity in the southeast of Yun-Lin County and the north of Chia-Yih County in Taiwan to reveal and map the spatial characteristics of transmissivity for hydrogeological study. The spatial maps of transmissivity were estimated and simulated using ordinary kriging, sequential Gaussian simulation and simulated annealing methods. Correlation analysis revealed that realizations of simulations could fully display the characteristics of transmissivity within this study area. The spatial maps of the estimation and simulation of transmissivity indicated that sequential Gaussian simulation and simulated annealing could not only reproduce the statistics and spatial variation of the measured transmissivity, but could also identify the global spatial continuity patterns of transmissivity in this study area. The realizations generated by sequential Gaussian simulation displayed significantly higher local heterogeneity than those generated by simulated annealing. The realizations of simulated annealing simulation are consistent rather than consistently in presenting the spatial patterns of transmissivity.

Keywords: Transmissivity; Simulated Annealing; Sequential Gaussian simulation, Ordinary kriging; Spatial characteristics

1. INTRODUCTION

Transmissivity and hydraulic conductivity are fundamental parameters of hydrogeological properties in characterizing aquifers for groundwater models. Proper modeling of preferential flow paths and of transport behavior requires the use of transmissivity fields that reproduce the spatial variability patterns observed in the field [Capilla et al., 1998]. These transmissivity fields sometimes contain significant uncertainties, including complex (unexplainable) variations in observed values of measurable attributes over the investigated area. These spatial and temporal variations can be extremely complicated. Thus, the reconstruction of the transmissivity field from the experimental hydraulic head data, an inverse problem, arises not only from the complexity of the diffusion equation linking the two variables, but also from considering the physical aspects of the site under study; such as the boundary conditions, the effective recharge, and the geology [Roth, 1998]. Therefore, it is important to characterize the spatial variabilities of transimissivity for groundwater models. Because of these variations in hydrogeological properties numerous authors have used statistical procedures to model spatial structures of interesting geohydrologic and physicochemical properties. Examples of such works include Eggleston at al. [1996], Fabbri [1997], Christensen [1997], Di Federico and Neuman [1997], Salandin and Fiorotto [1998], among others.

Geostatistical techniques like kriging incorporate the spatial or temporal characteristics of actual data into statistical estimation. If data appear highly continuous in space, points closer to the estimates receive higher weights than those farther away. Kriging estimates can be regarded as the most accurate linear estimator (i.e. Best Linear Unbiased Estimator). The estimated values based on kriging display a lower variation than the actual investigated values. To correct this shortcoming, geostatistical simulation can be performed. Simulation generates equally likely sets of values for a variable, which are

consistent with available in-situ measurements. This frequently implies that the simulated values have the same mean and variogram as the original data, and may also coincide with the original data at measurement points. Simulation focuses mainly on reproducing the fluctuations in the observations, instead of producing the optimal prediction [Sterk and Stein, 1997].

Geostatistical conditional simulation, such as by simulated annealing and sequential Gaussian simulation, attempts not only to generate a set of values with some specified mean and covariance, but also to reproduce observed data at several locations. Therefore, geologists use simulation to visualize fluctuations in major geologic patterns, investigate the fossil morphology, and map stratigraphical and structure surfaces [Christakos, 1992]. Recent works are Varljen and Shafer [1991], Eggleston et al. [1996], Mowrer [1997], Sterk and Stein [1997], Kentwell et al. [1999], Wang and Zhang [1999], Lin and Chang [2000a], Lin et al. [2000], and Lin and Chang [2000b].

This work uses conditional simulation techniques and ordinary kriging to produce the realizations and maps of trasmissivity in a real case study. The descriptive statistics, spatial structure (experimental variogram), correlation and spatial patterns of estimated and simulated results are also discussed herein. Finally, the estimation and simulation results are mapped into GIS to compare with the spatial distribution of geological formation for characterizing the spatial distribution of estimated and simulated transmissivity with measured data in the study area.

2. MATERIALS AND METHODS

In this study, the selected area (154.67 km^2) is on the east banks of the middle and upper streams of the Peikang River in the southeast part of Yun-Lin County and the north part of Gia-Yih County, Taiwan (Figure 1). The geological features of the study area from Dulliu Hill to the Peikang River include the Toukoshan Formation, Lichi Formation, Terrace Deposits and Alluvium, as presented in Figure 1.

The fan has a typical alluvia fan stratum structure, with a thick gravel layer on the east side reducing gradually to the west and southwest, while the muddy and sandy stratum thickens [Lin et al. 2000]. Tsao [1982] had studied and reported the field data of the well-drilling log and the aquifer of this study area. According to the reports of Tsao [1982] and Lin et al. [2000] Tapei city has the deepest well bores in the county, with an average depth of 199.7m. Meanwhile, wells in Talin Town ranks second, with an average depth of 179.3m, and wells in Tounan and Kukeng towns are 176.5m and 106.5m, respectively [Tsao 1982; Lin et al., 2000]. The average thickness of the aquifer in the well-drilling log ranges from 30.6-52.8m. The depth of the aquifer in the alluvia fan part of the Chuo-Shuei River and the West Bank of the Peikang River is 80m at Huwei town and an average of 100m at Tuku and Yuanchang towns [Lin et al. 2000]. Figures 2a and 2b present the locations of sampling wells of transmissivity data were measured by using the pumping test and provided by Yu-Ling Irrigation Association in. The measured values of transmissivity ranged from 8.637 to 407.483 (m^2/hr) at the sampling points of this study area.

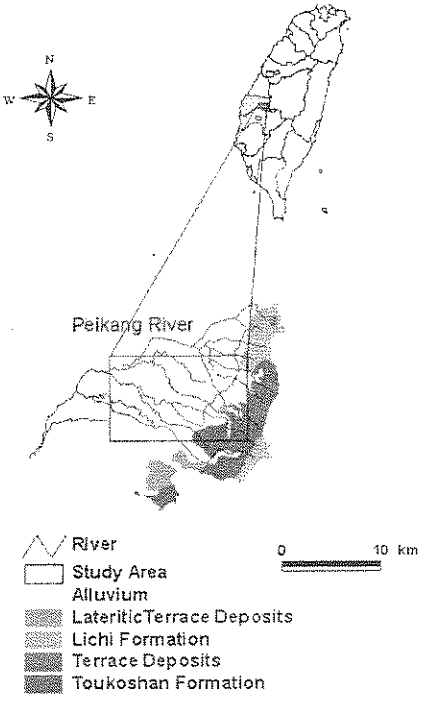

Figure 1. The location of study area

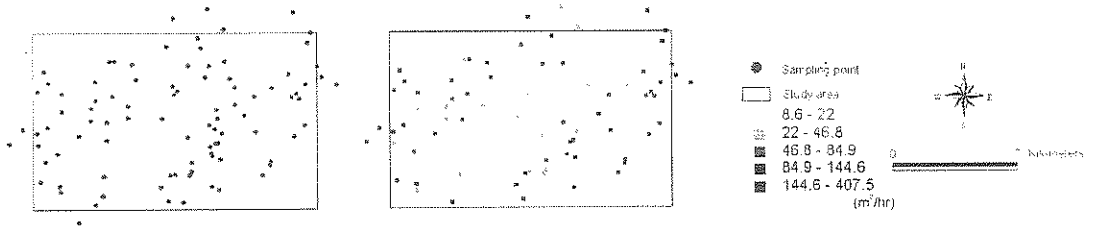

Figure 2. (a) The locations of sampling wells in study area; (b) Measured Values of transmissivity

Geostatistics provide a variogram of data within a statistical framework, including spatial and temporal covariance functions. As expected, these variogram models are termed spatial or temporal structures, and are defined in terms of the correlation between any two points separated either spatially or temporally.

Variograms provide a means of quantifying the commonly observed relationship whereby samples close together tend to have more similar values than samples farther apart. The variogram γ (h) is defined as:

$$\gamma(h) = (1/2)Var[Z(x) - Z(x+h)] \quad (1)$$

where h denotes the lag distance separating pairs of points, *Var* represents the variance of the argument, Z(x) is the value of the regionalized variable of interest at location x, and Z(x + h) denotes the value at the location x + h.

An experimental variogram γ (h), is given by:

$$\gamma(h) = 1/[2n(h)] \sum_{i=1}^{n(h)} [Z(x_i + h) - Z(x_i)]^2 \quad (2)$$

where γ (h) denotes the variogram for interval lag distance class h, and n(h) represents the number of pairs separated by the lag distance h.

Ordinary kriging, as applied within moving data neighborhoods, is a non-stationary algorithm corresponding to a non-stationary random function model with varying mean but stationary covariance [Deutsch and Journel, 1992]. Kriging estimates are weighted sums of the adjacent sample concentrations. The weights depend on the correlation structure exhibited. The criterion for selecting these weights is to minimize estimation variance. In this framework, kriging estimates can be regarded as the most accurate linear estimator (i.e. Best Linear Unbiased Estimator). At an unsampled location and for a given variogram, a kriging estimate can simply be considered an optimally weighted average of the surrounding sampled data [Cressie, 1990]. Kriging estimates the value of the random field at an unsampled location X_0 based on the given measured values in the linear form [Rouhani, 1985]

$$Z^*(X_0) = \sum^{N} \lambda_{i0} Z(X_i) \quad (3)$$

where $Z^*(X_0)$ denotes kriging estimates at X_0, $Z(X_i)$ represents measured values at X_i, $i = 1, \cdots, N$, and λ_{i0} is kriging weight for $Z(X_i)$ to estimate $Z^*(X_0)$.

2.2 Sequential Gaussian Simulation

Sequential Gaussian simulation assumes a Gaussian random field, thus the conditional cumulative density function (cdf) is completely characterized by the mean value and covariance [Fredericks and Newman, 1998]. In the sequential Gaussian simulation process, simulation is conducted upon the Gaussian transformation of the available measurements, so that each simulated value is conditional on the original data and all previously simulated values [Deutsch and Journel, 1992; Rouhani et al., 1995]. A simulated value at a one location is randomly selected from the normal distribution function defined by the kriging mean and variance based on neighborhood values. Finally, the simulated normal values are back transformed into simulated values for an original variable. The simulated value at the new randomly visited point value is dependent upon both the original data and previously simulated values. This process is repeated until all points are simulated.

2.3 Simulation by Simulated Annealing

The annealing algorithm requires that the image is perturbed by simulating thermal perturbation [Deutsch and Cockerham, 1994]. However, simulated annealing is an optimization technique to generate an initial field by drawing random values from a given histogram. Swapping the values in pairs of grid nodes not involving a conditioning datum, sequentially modifies this initial field. A swap is accepted if the objective function is lower [Deutsch and Journel, 1992]. This objective function (O) is defined as an average squared difference between the experimental and given variogram.

$$O = \sum_h \frac{[\gamma^*(h) - \gamma(h)]^2}{\gamma(h)^2} \quad (4)$$

Where γ(h) is the prespecified variogram, and γ^*(h) represents the variogram of the simulated realization.

A temperature function (the Boltzman distribution) in the simulated annealing procedure controls how the speed at which the optimization function is reduced by allowing certain switches that increase the optimization function [Deutsch and Journel, 1992; Eggleston et al., 1996]. The parameter, t, of the temperature function is termed the temperature in the annealing procedure. The higher the temperature, the more chance of an unfavorable swap being accepted [Deutsch and Journel, 1992].

$$P\{accept\} = \begin{cases} 1, & if O_{new} \le O_{old} \\ e^{\frac{O_{new} - O_{old}}{t}}, & otherwise \end{cases} \quad (5)$$

In this work the variogram models of transmissivity are also fitted within GS+ [Gamma Design Software, 1995]. The source codes of OKB2DM, SGSIM and SASIM in GSLIB [Deutsch and Journel, 1992] were modified to perform ordinary kriging and simulated annealing for estimating and simulating transmissivity. These simulations and estimates were performed in a square 38 column by 28 row grid

comprising 1664, 50m by 50m cells. Five simulations of sequential Gaussian and simulated annealing are performed in these 1664 cells. The results are transferred into Arcview 3.0a [ESRI, 1998] to display them and identify the spatial patterns of transmissivity.

3. RESULT AND DISCUSSION

3.1 Variography

A reasonably consistent set of best-fit models with minimum RSS (Model reduced sum of squares) and maximum r^2 (Regression coefficient) values were generated by least squares model fitting of variogram of transmissivity data.

3.2 Statistics Concerning Estimation and Simulation

The ordinary kriging estimates and conditional simulations were based on the above variogram models and the 92 transmissivity observations. Tables 1 summarize the descriptive statistics related to ordinary kriging, Gaussian sequential simulation and simulated annealing results. The mean and standard deviation values of kriging estimation of transmissivity are 87.066 (m^2/hr) and 39.848 (m^2/hr). Meanwhile, the median values [61.611-72.950(m^2/hr) and 62.265-70.000(m^2/hr)] of the transmissivity simulations using sequential Gaussian and simulated annealing were almost identical to that [64.275(m^2/hr)] of the measured transmissivity data.

Moreover, the standard deviation values [85.074-91.010(m^2/hr) and 87.643-94.618(m^2/hr)] of the results from sequential Gaussian and simulated annealing simulation also approached the empirical transmissivity data value [92.082(mg/kg)].

Gaussian simulated realizations have a common origin and exhibit a similar spatial pattern of tranmissivity in the study area.

Table 3 lists Pearson correlation coefficients among kriging estimated results and simulated realizations of simulated annealing. The range of correlation coefficients between kriging estimates and each realization varies from 0.577 to 0.787, and exhibits a strong linear correlation, as illustrated in Table 3.

Table 1. Descriptive statistics of kriging estimates and simulations

	N	Mean	Median	Std. Dev.	Skewness	Kurtosis	Min	Max
OK	1664	87.066	73.366	39.848	1.211	0.330	33.783	196.713
SGS1	1664	90.203	61.611	88.642	1.915	3.557	8.709	407.971
SGS2	1664	91.790	63.849	90.753	1.865	3.222	8.650	407.980
SGS3	1664	99.109	72.950	94.618	1.761	2.695	8.697	407.859
SGS4	1664	91.165	64.274	89.028	1.908	3.469	8.642	407.920
SGS5	1664	86.736	59.406	87.643	1.917	3.494	8.649	407.690
SAS1	1664	90.433	64.749	85.074	1.866	3.577	8.637	407.483
SAS2	1664	93.587	65.583	91.010	1.803	3.027	8.637	407.483
SAS3	1664	94.880	70.000	90.093	1.771	2.963	8.637	407.483
SAS4	1664	90.290	62.265	89.262	1.900	3.427	8.637	407.483
SAS5	1664	91.251	65.593	87.826	1.868	3.423	8.637	407.483

OK: Ordinary kriging; SGS: Sequential Gaussian simulation; SAS: Simulated annealing simulation.

Table 2. Pearson correlation coefficients of estimation and sequential Gaussian simulation

	OK	SGS1	SGS2	SGS3	SGS4	SGS5
OK	1.000	0.209**	0.326**	0.175**	0.215**	0.171**
SGS1	0.209**	1.000	0.135**	0.120**	0.096**	0.114**
SGS2	0.326**	0.135**	1.000	0.124**	0.093**	0.141**
SGS3	0.175**	0.120**	0.124**	1.000	0.069**	0.054*
SGS4	0.215**	0.096**	0.093**	0.069**	1.000	0.059*
SGS5	0.171**	0.114**	0.141**	0.054*	0.059*	1.000

* Correlation is significant at the 0.05 level (2-tailed test)
** Correlation is significant at the 0.01 level (2-tailed test).
OK: Ordinary kriging;
SGS: Sequential Gaussian simulation

Table 3. Pearson correlation coefficients of estimation and simulated annealing simulation

	OK	SAS1	SAS2	SAS3	SAS4	SAS5
OK	1.000	0.773**	0.577**	0.656**	0.787**	0.685**
SAS1	0.773**	1.000	0.602**	0.712**	0.842**	0.739**
SAS2	0.577**	0.602**	1.000	0.611**	0.649**	0.688**
SAS3	0.656**	0.712**	0.611**	1.000	0.629**	0.624**
SAS4	0.787**	0.842**	0.649**	0.629**	1.000	0.811**
SAS5	0.685**	0.739**	0.688**	0.624**	0.811**	1.000

* Correlation is significant at the 0.05 level (2-tailed test)
** Correlation is significant at the 0.01 level (2-tailed test).
OK: Ordinary kriging
SAS: Simulated annealing simulation.

These results also imply that the kriging estimates

and simulated annealing simulated realizations have a common origin and exhibit a similar spatial pattern of tranmissivity in the study area. Moreover, the correlation coefficients among the realizations are strongly significant at the 0.01 probability level according to the 2-tailed test, as presented in Table 3. The range of these coefficients is from 0.602 to 0.842.

3.4 Spatial Patterns of Transmissivity

The estimates and simulations were mapped using Arcview 3.0a (Figures 4 and 5). The ordinary kriging estimate maps (Figure 4a) confirmed that kriging tended to show spatial patterns but smooth extreme values of the empirical transmissiivity data set. These kriging results might overestimate the size of high and low concentration areas of transmissivity, and underestimate areas with extremely high and low concentrations of transmissivity.

The kriging estimated and conditional simulated transmissivity maps illustrate that the low values formed an approximately triangular shape in the center of this study area (Figures 4 and 5). These maps display that the transmissivity values gradually increased from this triangular shape area to the surrounding area of this center area, as illustrated in Figures 4 and 5. The high transmissivity values were located on the Dulliu Hill in the east part of the study area and on the west of the Peikang River in the northwestern part of the study area, as illustrated in Figures 4 and 5. Moreover, the transmissivity values of the study area gradually decreased from the Dulliu Hill to the center of the area.

The above-simulated maps of transmissivity also illustrate a high heterogeneity area of transmissivity especially in sequential Gaussian simulated maps. These simulated maps emphasize the significant variation across short distances and provide a measure of spatial uncertainty. These maps also reveal that the realizations of transmissivity generated by sequential Gaussian simulation reflected more local heterogeneity than those generating by simulated annealing. However, the simulated annealing simulated transmissivity maps fully display the characteristics of the geological formation of this study area (Figures 1, 4 and 5). These maps also confirm that the study area has a rather typical alluvia fan stratum structure, since the gravel layer on the east side is thick, gradually thinning towards the west and south-west, while the thickness of the muddy and sandy stratum increases correspondingly.

4. CONCLUSIONS

This work has demonstrated that kriging and conditional simulation techniques, including Gaussian sequential simulation and simulated annealing can be used to identify pollution sources and patterns, ordinary kriging failed to reproduce the statistics of transmissivity better than simulation techniques. In addition to reproducing the spatial variation of the measured transmissivity, sequential Gaussian simulation and simulated annealing also simulated global spatial continuity and discontinuity patterns. Moreover, the simulated annealing method achieved more accurate results than sequential Gaussian simulation and kriging by comparing to global statistics and spatial patterns of measured transmissivity data. Finally, the realizations of simulated annealing simulation are consistently in presenting the spatial patterns of transmissivity in this study area.

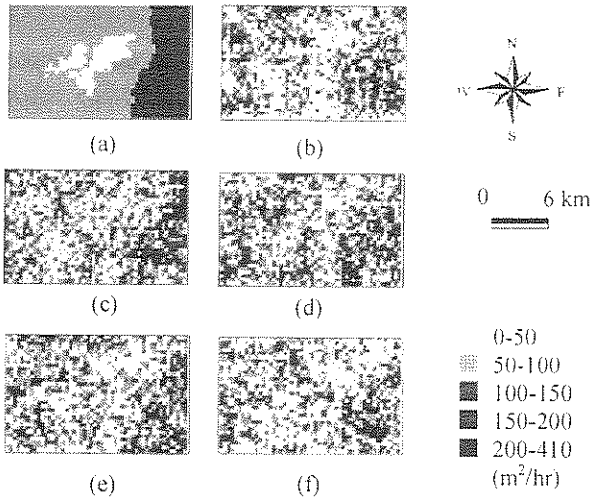

Figure 4. The spatial maps of Transmissivity of (a) kriged value; (b) SGS1; (c) SGS2; (d) SGS3; (e) SGS4; (f) SGS5

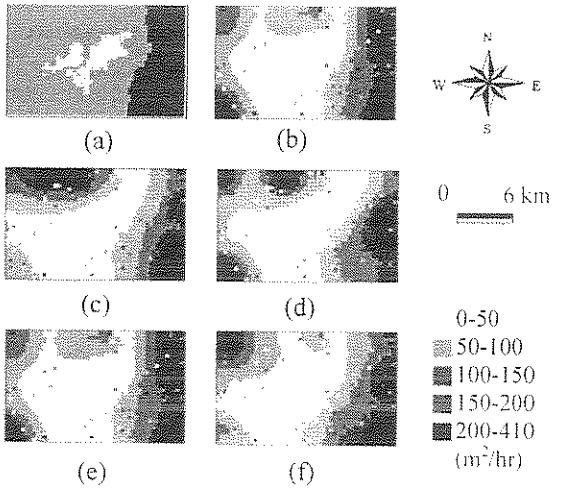

Figure 5. The spatial maps of Transmissivity of (a) kriged value; (b) SAS1; (c) SAS2; (d) SAS3; (e) SAS4; (f) SAS5.

5. REFERENCE

Capilla, J.E., J. Rodrigo and J.J. Gomez-Hernandez, Worth of secondary data compared to piezometric data for the probabilistic assessment of radionuclide migration, *Stochastic Hydrology and Hydraulic*, 12, 191-204, 1998.

Christakos, G., Random field models in earth sciences, Academic Press, Inc, New York, 1992.

Christensen, S., On the strategy of estimating regional-scale transmissivity fields, *Ground Water*, 35(1), 131-139, 1997.

Cressie, C., The origins of kriging, *Math. Geology* 22(2), 239-252, 1990.

Deutsch, C.V. and P.W. Cockerham, Practical considerations in the application of simulated annealing of stochastic simulation. Math. Geology 26(1), 67-82, 1994

Deutsch, C.V. and A.G Journel., GSLIB, Geostatistical software library and user's guide. Oxford University Press, New York, 1992.

Di Federico V. and S.P. Neuman, Scaling of random-fields by means of truncated power variograms and associated spectra, *Water Resour. Res.*, 33(5), 1075-1085, 1997.

Eggleston, J.R., S.A. Rjostaczer, and J.J. Peirce, Identification of hydraulic conductivity structure in sand and gravel Aquifers: Cape cod data set, *Water Resour. Res.*, 32(5), 1209-1222, 1996.

ESRI (Environmental systems Research Institute) Arcview 3.0; ESRI, Redlands, CA, USA, 1998.

Fabbri, P., Transmissivity in the geothermal Euganean basin: A geostatistical analysis, *Ground Water*, 35(5), 881-887, 1997.

Fredericks, A.K., K.B. Newman, A comparison of the sequential Gaussian and Markov-Bayes simulation methods for small samples. Math. Geology 30(8), 1011-1032, 1998

Gamma design software, GS+: Geostatistics for the environmental sciences. Version 2.3. Gamma Design Software, Plainwell, MI, 1995.

Kentwell, D.J., L.M. Bloom and G.A.Comber, Improvements in grade tonnage curve prediction via sequential Gaussian fractal simulation. *Math. Geology*, 31(3), 311-325, 1999.

Lin, Y.P. and T.K. Chang, Geostatistical estimation and simulation of the spatial variability of soil zinc, *Journal of Environ. Sci. Health Part A*, 35(3), 327-347, 2000a.

Lin, Y.P. and T.K. Chang, Simulated annealing and kriging method for identifying the spatial patterns and variability of soil heavy metal, *Journal of Environ. Sci. Health Part A*, 35(7), 1089-1115, 2000b.

Lin, Y.P., C.C. Lee and Y.C.Tan, Geostatistical approach for identification of transmissivity structure at Duliu area in Taiwan. *Environ. Geol.*, 40(1-2), 111-120, 2000.

Mowrer, H.T., Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modelling*, 98, 73-86, 1997

Roth, C., Is lognormal kriging suitable for local estimation? *Math. Geol.* 30(8), 999-1009. 1998.

Rouhani, S., Variance reduction analysis, Water Resour. Res., 21(6), 837-864, 1985.

Rouhani, S., Y.P. Lin and Y. Shi, H-area/ITP geostatistical assessment in-situ and engineering properties, final technical report, Westinghouse Savannah River Company, 1995.

Salandin, P. and V. Fiorotoo, Solute transport in highly heterogeneous aquifers, *Water Resour Res.*, 34(5), 949-961, 1998.

Sterk, G. and A. Stein, Mapping wind-blown mass transport by modeling variability in space and time, *Soil Sci. Soc. Am. J.*, 61, 232-239, 1997.

Tsao, E.S., The study of groundwater mathematical model in Yu-Ling County, Agricultural Engineering Center, Tau-Yuan, Taiwan, 1982.

Varljen, M.D. and Shafer J.M., Assessment of uncertaity in timerelated captures zones using conditional simulation of hydraulic conductivity, *Ground Water*, 29, 737-748, 1991.

Wang, X.J. and Z.P. Zhang, A comparision of conditional simulation, kriging and trend surface analysis for soil heavy metal pollution pattern analysis, *J. Environ. Sci. Health Part A*, 34(1), 73-89, 1999.