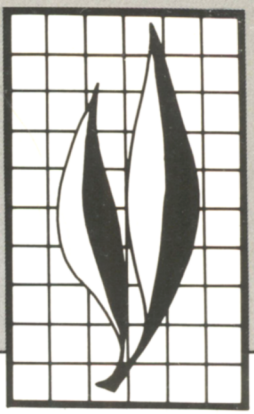
Hilgardia

A JOURNAL OF AGRICULTURAL SCIENCE PUBLISHED BY
THE CALIFORNIA AGRICULTURAL EXPERIMENT STATION

Volume 55 • Number 4 • July 1987



The Spatial Variability of Water and Solute Transport Properties in Unsaturated Soil

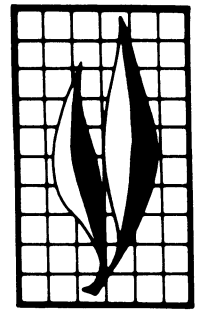
I. Analysis of Property Variation and Spatial Structure with Statistical Models

William A. Jury, David Russo, Garrison Sposito, and Hesham Elabd

II. Scaling Models of Water Transport

William A. Jury, David Russo, and Garrison Sposito

UNIVERSITY OF CALIFORNIA DIVISION OF AGRICULTURE AND NATURAL RESOURCES



ABSTRACTS

I. Analysis of Property Variation and Spatial Structure with Statistical Models

This review presents and examines relevant information from existing spatial variability studies of soil water and solute transport properties. Although most of the information available allowed only a conventional statistical analysis (mean and variance) of the pertinent properties, the field studies of Nielsen, Biggar, and Erh (1973) and Russo and Bresler (1981) were also suitable for spatial structure analysis. Detailed structural analysis of the saturated hydraulic conductivity (K_s) of these two fields demonstrated how this type of analysis may reveal field characteristics that are not apparent from conventional statistical analysis.

Using the Akaike Information Criterion for model discrimination, the three-dimensional spatial distributions of $\ln K_s$ of both fields were shown to be described best by a spherical covariance function and a linear drift function. The Hamra field of Russo and Bresler (1981) had a much larger deterministic drift component and a smaller stochastic

Continued inside back cover.

THE AUTHORS:

William A. Jury is Professor of Soil Science and Soil Physicist, Department of Soil and Environmental Sciences, University of California, Riverside.

David Russo was Visiting Scientist in the Department of Soil and Environmental Sciences, University of California, Riverside, from 1984 to 1986. He has since returned to his position as Soil Physicist at the Volcani Center, Agricultural Research Organization, Bet-Dagan, Israel.

Garrison Sposito is Professor of Soil Science and Soil Chemist, Department of Soil and Environmental Sciences, University of California, Riverside.

Hesham Elabd, former Graduate Research Assistant in the Department of Soil and Environmental Sciences, University of California, Riverside, is now Postdoctoral Researcher at New Mexico State University, Las Cruces.

William A. Jury, David Russo, Garrison Sposito, and Hesham Elabd

The Spatial Variability of Water and Solute Transport Properties in Unsaturated Soil[1]

I. Analysis of Property Variation and Spatial Structure with Statistical Models

INTRODUCTION

THE INHERENT TEXTURAL and structural variabilities in field soils and hydrogeologic formations are widely recognized as dominant factors influencing mass transport through the subsurface zone (Beckett and Webster 1971; Freeze 1975; Hoeksema and Kitanidis 1985; Warrick and Nielsen 1980). Because of this spatial variability, the transport and retention properties of a field-scale soil or aquifer unit also exhibit spatial variation in response to such inputs as irrigation, rain, natural or artificial recharge, and discharge from waste disposal sites. Pronounced lateral and vertical variations in the water or chemical transport properties of the vadose and groundwater zones seriously limit the applicability of the traditional, column-scale deterministic approaches generally used to analyze water and solute flow processes (Bear 1972; Nielsen, Biggar, and Erh 1973). This problem has led to the development of stochastic models to describe solute movement in groundwater (Dagan 1982, 1984; Gelhar and Axness 1983; Gelhar, Gutjahr, and Naff 1979; Smith and Schwartz 1980) and in the unsaturated zone (Amoozegar-Fard, Nielsen, and Warrick 1982; Bresler and Dagan 1979, 1981, 1983; Dagan and Bresler 1979, 1983; Jury 1982; Russo 1984b; Simmons 1982). Although the underlying physics is similar, groundwater transport and transport through the unsaturated zone do differ with respect to scale, flow regime, and the direction of the principal velocity component relative to the direction of the principal variation in porous medium properties.

Studies describing groundwater transport deal with a system in which the vertical dimension is relatively small in scale compared with the horizontal dimension, which may be on the order of a kilometer or more. Because the medium is saturated, water-flow characterization requires only a knowledge of the saturated hydraulic conductivity (or transmissivity) and the storage capacity of the aquifer; solute-flow characterization requires a knowledge of the components of the dispersion tensor as well. Often, researchers will assume that this tensor reduces to three principal components representing longitudinal and transverse dispersion. A common approach in modeling water and solute transport through groundwater is to average the transport properties

[1]Accepted for publication January 29, 1987.

over depth, producing an equivalent system viewed as two-dimensional in the horizontal plane. Given that the water flux is horizontal, groundwater transport takes place in the direction of heterogeneity in the medium.

Studies describing transport through unsaturated soil are also concerned with a system in which the vertical plane is small in scale compared with the horizontal plane that defines a field area, which may be measured in the tens or hundreds of meters. In most cases, water input to the soil is time dependent and the soil remains unsaturated, so transport and retention functions that depend on the water content are used to characterize the soil; solute flow characterization requires in addition a knowledge of the components of the dispersion tensor, which also will depend on the degree of saturation of the medium. A common approach employed in describing field-scale water and solute movement is to view the transport process locally as a vertical, one-dimensional flow. According to this perspective, the system is visualized as a collection of one-dimensional, noninteracting soil columns having different properties (Dagan and Bresler 1979). In contrast to the situation in groundwater, a transport process through the vadose zone usually takes place in a direction perpendicular to any layering in the medium.

Stochastic transport models share the common assumption that the porous medium may be considered as a single realization of a random field. The application of these models, therefore, relies on a quantification of the spatial variability of the random field that is assumed to represent the pertinent transport properties. A complete characterization of the random field for a given property would require specification of the joint probability density function (PDF) for the property at established coordinate points throughout the system. This, unfortunately, is generally impossible. An alternative approach involves the quantification of field spatial variability in terms of the first and second statistical moments. Even this approach, however, has formidable data requirements. In the last decade only a few experimental studies have been designed to characterize the spatial variability of soil water and solute transport properties. With respect to groundwater hydrology, the studies of Freeze (1975) and Delhomme (1979), and the recent extensive survey of various aquifers by Hoeksema and Kitanidis (1985) give the most comprehensive information available on the spatial variability of aquifer transmissivity and storage capacity; limited information on the spatial variability of longitudinal dispersivity is summarized by Gelhar (1986). With respect to the vadose zone, pioneering work was done by Nielsen, Biggar, and Erh (1973) on spatial variations in the hydraulic conductivity and water retention functions, followed by the study of Biggar and Nielsen (1976) examining spatial variability in the solute velocity and longitudinal dispersion coefficient. Other studies that give information about spatial variation in the hydraulic conductivity or soil water retention functions are those of Babalola (1978), Cameron (1978), Russo and Bresler (1981), Jones and Wagenet (1984), Russo (1983, 1984b), and Greminger, Sud, and Nielsen (1985). A recent study by Russo (1984b) has provided additional information on the spatial distribution of the longitudinal dispersivity.

The objectives of the present study are first to examine existing spatial variability studies of soil water and solute transport properties with respect to trends in variation, and second to review experimental evidence in support of the different statistical models used to characterize spatial variability.

CONCEPTS AND DEFINITIONS

Basic Assumptions

The porous medium is regarded as a macroscopic continuum whose properties are continuous functions of the space coordinates. The point-value of a property $u(x)$ is viewed as the value measured on an element of the medium centered at the point x. Because $u(x)$ may vary in a highly irregular fashion in space, a set of its measured values is interpreted as a realization $U(x, \omega_1)$ of a spatial stochastic function $U(x, \omega)$ for which the probability distribution function has to be inferred. Here ω is a parameter denoting a random function whose values (ω_1, etc.) denote sample realizations (Doob 1953).

Strictly, $U(x, \omega)$ should be regarded as a three-dimensional, anisotropic stochastic function. The paucity of experimental data available from a given field site, however, usually limits the detection of spatial anisotropies. In most unsaturated-zone studies, the horizontal plane is much larger in scale than the vertical plane, and it is therefore common to neglect property variations with depth in favor of depth-averaged values, so that $U(x, \omega)$ is interpreted as a two-dimensional, isotropic stochastic function in the horizontal plane (Russo and Bresler 1981).

The Ergodic Hypothesis

Conceptually, $U(x, \omega)$ may be considered an ensemble of realizations that share the same statistical properties. The ensemble concept is convenient for defining the statistical properties of $U(x, \omega)$. Physically, the ensemble mean can be understood as the arithmetic average of repeated measurements of a property at a given spatial point, under the same external conditions. In practice, however, only one realization of $U(x, \omega)$ will be available and the ergodic hypothesis (Lumley and Panosfky 1964; Sposito, Jury, and Gupta 1986) must be invoked. This hypothesis states that inferences about the statistical structure of $U(x, \omega)$ may be based on estimates of the ensemble averages gained from spatial averages obtained from a single realization of $U(x, \omega)$.

Second-Order Stationarity

A complete statistical description of $U(x, \omega)$ requires specification of the joint PDF at all points in space. To simplify matters in compliance with available data sets, the quantification of $U(x, \omega)$ is restricted to its first and second statistical moments. Furthermore, it is assumed that $U(x, \omega)$ can be described by the general model (Delfiner 1976)

$$U(x, \omega) = M(x) + Z(x, \omega), \qquad [1]$$

where $M(x)$ is a prior mean or drift function, and $Z(x, \omega)$is a zero-mean stochastic function characterized completely by a covariance function, C_Z. Usually $Z(x, \omega)$ is assumed to be a second-order stationary random function, which means that its first

two statistical moments are invariant under spatial translation (Doob 1953). The expectation of $Z(x,\omega)$ is thus

$$E[Z(x,\omega)] = 0, \quad [2]$$

and the covariance of $Z(x,\omega)$ between any two points, x_1 and x_2, does not depend individually on x_1 and x_2 but only on the lag vector, $h = x_1 - x_2$:

$$E\{[Z(x,\omega)]\,[Z(x+h,\omega)]\} = C_Z(h). \quad [3]$$

In particular, the variance of $Z(x,\omega)$ is independent of x:

$$E[Z(x,\omega)^2] \equiv \text{var}[Z(x,\omega)] = C_Z(0). \quad [4]$$

In the general case described by equation 1, the first moment of $U(x,\omega)$ is $E[\mathrm{U}(x,\omega)] = M(x)$, and the second central moment is $E[U(x+h,\omega)\,U(x,\omega)] - M(x)\mathrm{M}(x+h) = C_U(h,x)$. When drift is present, the first two moments of $U(x,\omega)$ are functions of x, hence $U(x,\omega)$is not second-order stationary. When $M(x)$ is a constant (denoted by m), however, $E[U(x,\omega)] = m$, $C_U(h) = C_Z(h) \equiv C(h)$, and $U(x,\omega)$ is second-order stationary. For a second-order stationary function, two equivalent functions for characterizing the spatial structure of $U(x,\omega)$ may be defined as

$$\frac{E\{[U(x,\omega)-m]\,[U(x+h,\omega)-m]\}}{E\{[U(x,\omega)-\mathrm{m}]^2\}} = \frac{C(h)}{C(0)} = \rho(h), \quad [5]$$

where $\rho(h)$ is the correlogram or autocorrelation function, and

$$\frac{E\{[U(x+h,\omega)-U(x,\omega)]^2\}}{2} = C(0) - C(h) = \gamma(h), \quad [6]$$

where $\gamma(h)$ is the semivariogram.

The Intrinsic Hypothesis

Instead of assuming second-order stationarity, we can assume a weaker hypothesis, the intrinsic hypothesis (Delhomme 1978), which requires stationarity only for increments of $U(x,\omega)$. For an intrinsic stochastic function $U(x,\omega)$, the increment $[U(x+h,\omega) - U(x,\omega)]$ has zero expectation,

$$E[U(x+h,\omega) - U(x,\omega)] = 0, \quad [7]$$

and a uniform variance,

$$E\{[U(x+h,\omega) - U(x,\omega)]^2\} = 2\,\gamma(h). \quad [8]$$

Second-order stationarity implies the intrinsic hypothesis, but the intrinsic hypothesis does not imply second-order stationarity.

Correlation Scale

When $U(x,\omega)$ is a second-order stationary function, we can define a characteristic length (correlation) scale in terms of the integral scale (Russo and Bresler 1981):

$$J = \int_0^\infty \rho(h)dh. \qquad [9]$$

For a second-order stationary function, $\rho \to 0$ as $h \to \infty$, and the integral scale can be interpreted as the largest average distance over which $U(x_1,\omega)$ and $U(x_2,\omega)$ correlate.

The semivariogram γ and the autocorrelation function ρ are cornerstones of geostatistical interpolation and estimation procedures (Journel and Huijbregts 1978) and stochastic continuum transport modeling (Bakr et al. 1978; Gelhar and Axness 1983), but the estimation of these functions along with their associated correlation scales may be very complicated. Briefly, three main sources of error may affect the estimates: (1) only one realization of the property is available from which all statistical estimates of the population must be inferred; (2) the number and location of the sampling points may not be sufficient to encompass the scale of the field variability, or to make an accurate determination of the spatial structure; and (3) the presence of an undetected drift component may affect semivariogram estimates.

With respect to item 1, the statistical inference of any of the properties of $U(x,\omega)$ is affected by the finite size of the available realization. Russo and Bresler (1982) used the ergodic hypothesis to derive relationships of the error in estimating the ensemble mean of $U(x,\omega)$ that is caused by finite size to the size of the field area that should be sampled for two-dimensional stationary functions.

With respect to item 2, most experienced geostatisticans would concede that many samples (100 or more) are required for the accurate estimation of the semivariogram. In practice, however, this requirement has been difficult to satisfy, especially for transport properties, which are difficult to measure. Bresler and Green (1982) and Russo (1984a) considered the case of a relatively small sample size and formulated the problem of semivariogram estimation in terms of the design of the sampling network. Cressie and Hawkins (1980), Cressie (1984), and Dowd (1984) have introduced robust and resistant methods for the estimation of the semivariogram. Estimation methods that involve a selection of parametric models for the covariance and drift functions and the estimation of their properties were introduced by Kitanidis (1983). Usually, a successful application of these methods requires that the data (or a transformation of data) follow a joint Gaussian distribution, and that the sample size be large enough for application of the asymptotic theory.

With respect to item 3, the presence of drift produces a nonstationary, random field, and any attempt to estimate the semivariogram, the correlogram, or the associated correlation scale requires the prior estimation of $M(x)$. In principle, parametric estimation methods (Kitanidis 1983) may be used to estimate the parameters of both $M(x)$ and $C_Z(h)$. Some of the spatial variability of $Z(x,\omega)$ may be fitted by the model used for $M(x)$, but this may produce considerable overestimation of the parameters of $C_Z(h)$ (Russo and Jury 1987b).

Because stochastic processes (Panchev 1971) or regionalized variables (Matheron 1971) have been introduced only recently to soil physics and groundwater hydrology, many attempts to quantify the spatial variability of a porous medium have been based on a conventional statistical approach (e.g., Cameron 1978; Freeze 1975; Jones and Wagenet 1984; Nielsen, Biggar, and Erh 1973; Rogowski 1972). In this approach, the observations of a given property are assumed to be statistically independent, regardless

of their spatial positions, and are used to estimate a PDF for the property. Subsequent statistical analyses such as linear regression (Nielsen, Biggar, and Erh 1973) or an analysis of the relationships between the error of estimation and the sample size (Keisling et al. 1977) also require that the observations be distributed normally.

For most natural porous media, spatial variations are not completely disordered but instead possess a structure of a typical size characterized by an integral scale, J. By neglecting this structure, researchers may come up with an incomplete statistical description of the medium and biased estimates of the PDF for the pertinent property, since no account is taken of the correlation between neighboring observations. On the one hand, unbiased estimates of the parameters of the PDF may be obtained if the distance between any pair of observation points is arranged to be much larger than the integral scale of the observed property. On the other hand, such a sampling network design will not reveal the correlation structure of the property.

The Lognormal Distribution

Transport properties measured near the soil surface have often been found to be skew-distributed—their sample frequency distributions are not well represented by a Gaussian distribution. Although the distribution of a given property cannot be inferred solely on theoretical grounds, the lognormal distribution (Aitcheson and Brown 1976; Hald 1952)

$$f(u) = \exp[-(\ln u - \mu)^2/2\sigma^2]/[(2\pi)^{1/2}\sigma u] \qquad [10]$$

has been found to describe adequately the distribution of skewed properties. In general, the lognormal distribution will provide a representation superior to the normal distribution for properties that (1) are positive definite ($u \geq 0$), (2) have a large coefficient of variation (CV), and (3) have a positively skewed PDF (i.e., the sample mean is larger than the mode and median). In addition, as the sample CV approaches the value 0.1, the lognormal distribution approaches symmetry and can adequately represent the normal distribution.

There are several advantages to using a lognormal distribution to represent the sample frequency. First, the parameter σ^2 describing the variance of the log-transformed property becomes a useful index of variability with which to compare different parameters and to study the influence of different characteristics like soil texture on the variability of a given property. Second, the lognormal distribution has the property that, if u is lognormally distributed with $E[\ln u] = \mu$ and $\text{var}[\ln u] = \sigma^2$, then $y = au^b$ also is lognormally distributed with $E[\ln y] = \ln a + b\mu$ and $\text{var}[\ln y] = b^2\sigma^2$ (Aitcheson and Brown 1976). This property will allow a variety of different experimental results to be analyzed in a common manner. For example, if an experiment showed a lognormal distribution of solute travel times from the soil surface to a depth L of observation, then it may be inferred from the relation $V = L/t$ between the travel time and the vertical velocity that $\text{var}[\ln t] = \text{var}[\ln V]$ and that V is also lognormally distributed.

Transport properties measured in the field have been found to be described better by a lognormal distribution than by a normal distribution (Biggar and Nielsen 1976; Freeze 1975; Hoeksema and Kitanidis 1985; Jury, Stolzy, and Shouse 1982; Nielsen, Biggar, and Erh 1973). The analytical techniques for parameter estimation applied in such

studies, however, except the study of Hoeksema and Kitanidis (1985), did not take into account correlations between neighboring observations. Hoeksema and Kitanidis gave a procedure for transforming correlated measurement values into correlated residuals, which in turn were transformed into uncorrelated residuals by using the estimated covariance function of $Z(x,\omega)$ and a transformation matrix proposed by Kitanidis (1983). Standard statistical methods (Haan 1977) may be used to analyze the distribution of these uncorrelated residuals and to validate the assumed statistical model for the spatial variability of a soil property.

ANALYSIS OF THE SPATIAL DISTRIBUTION OF TRANSPORT PROPERTIES

Water Transport Properties

Field studies of a given water transport property are difficult to compare because methodology frequently differs from one investigation to another and transport properties are usually estimated indirectly by assuming the validity of a specific physical model for the transport process (e.g., the one-dimensional Richards equation). Furthermore, transport functions such as unsaturated hydraulic conductivity are sometimes assumed to have a specific mathematical form (e.g., exponential) that may not be accurate for all soils. Finally, the number of samples used and the volume of soil sensed by a given sampling device may influence the measured variability, making comparisons between different studies difficult (Wagenet 1985).

The data in table 1 summarize the mean and coefficient of variation of the saturated hydraulic conductivity as measured in vadose-zone field studies, together with the parameter $\sigma \ln K_s$, the standard deviation of $\ln K_s$. The first three studies mentioned contain data from several depths as well as from different sites, so part of the data may be highly correlated and for that reason the statistical moments may not be accurate.

Although the studies in table 1 cover a variety of soil textures, no apparent relationship is present between the standard deviation parameter $\sigma_{\ln K_s}$ and soil type. The values of σ range between 0.5 and 1.6, and in only two of the studies did values exceed 1.2. The highest values (1.6 and 1.4) represented the variability among replicates of a series of coarse-textured soils (Willardson and Hurst 1965) and a large 150 ha field that ranged in texture from a loam to a silty clay loam (Nielsen, Biggar, and Erh 1973). With the exception of these studies, the values of σ span a relatively narrow range, with no apparent dependence on field size or sampling depth.

Table 2 contains a summary of studies on surface infiltration rates i measured in the field. With the exception of the study by Nielsen, Biggar, and Erh (1973), in which water was ponded to produce steady-state infiltration over 20 different large plot areas (6.2 × 6.2 m) located over 150 ha, all studies mentioned in our report sampled only small volumes of soil, either with ring infiltrometers or with inverse auger holes. For the most part, the standard deviation parameter $\sigma_{\ln i}$ is smaller than $\sigma_{\ln K_s}$ (table 1), both under steady-state and transient conditions.

The three values of infiltration rate parameters (table 2) reported by Sisson and Wierenga (1981) are notable because they represent measurements of i on the same 6.2 × 6.2 m field plot as determined with infiltrometer rings of differing sizes. For

TABLE 1. FIELD STUDIES OF THE SPATIAL VARIABILITY OF SATURATED HYDRAULIC CONDUCTIVITY

Mean	$\sigma_{\ln K_s}$	CV*	Texture	Area	Depth	n	Method of measurement	Reference
cm		%		*ha*	*cm*			
21	1.4	243	clay loam	150	0-150	120	ponded 6.2 × 6.2 m plots	Nielsen, Biggar, and Erh 1973
168	1.2	190	sandy loam	15	30-150	64	undisturbed ponded lab cores	Gumaa 1978
316	0.6†	69	sand	0.8	0-120	120	in situ permeameter	Russo and Bresler 1981
84	0.6	89	loamy sand	0.4	0-20	12	undisturbed ponded lab cores	Cassel 1983
4	0.5	48	silty clay loam	‡	30-90	33	undisturbed ponded lab cores	Willardson and Hurst 1965
19	0.9	103	coarse	‡	30-90	330	undisturbed ponded lab cores	Willardson and Hurst 1965
11	0.9	118	fine	‡	30-90	287	undisturbed ponded lab cores	Willardson and Hurst 1965
7	0.8	92	silty clay	‡	30-90	339	undisturbed ponded lab cores	Willardson and Hurst 1965
28	1.6	320	very coarse	‡	30-90	36	undisturbed ponded lab cores	Willardson and Hurst 1965
56	0.9	118	coarse	‡	30-90	352	undisturbed ponded lab cores	Willardson and Hurst 1965
71	0.9	105	loamy sand	‡	30-90	121	undisturbed ponded lab cores	Willardson and Hurst 1965
99	0.7†	81	loamy sand	91.6	0-15	5	undisturbed ponded lab cores	Babalola 1978
24	1.2†	178	silty loam	91.6	120-135	5	undisturbed ponded lab cores	Babalola 1978

*Coefficient of variation.

†Estimated from equation 13 and reported coefficient of variation.

‡Composite single soil series data from SCS, Imperial County.

TABLE 2. FIELD STUDIES OF THE SPATIAL VARIABILITY OF INFILTRATION RATE

Mean	$\sigma_{\ln K_s}$	CV*	Texture	Area	n	Method of measurement	Reference
cm		%		*ha*			
203	0.47	50	silty loam	9.6	26	double ring 0-30 minutes	Sharma, Gander, and Hunt 1980
15	1.18	94	clay loam	150	20	steady ponded 6.2 × 6.2 m plots	Nielsen, Biggar, and Erh 1973
17	0.38	40	loam	0.9	1,280	double ring steady state	Vieira et al. 1981
7	0.64	71	silty clay loam	0.004	625	5 transects of 5 cm diam. rings	Sisson and Wierenga 1981
9	0.52	56	silty clay loam	0.004	125	5 transects of 25 cm diam. rings	Sisson and Wierenga 1981
9	0.23	23	silty clay loam	0.004	25	5 transects of 125 cm diam. rings	Sisson and Wierenga 1981
47	0.78	79	7 series	100	20	double ring	Duffy, Wierenga, and Kselik 1981
263	0.84	97	7 series	100	15	inverse auger hole	Duffy, Wierenga, and Kselik 1981

*Coefficient of variation.

example, one study consisted of five rows containing 125 adjacent 5 cm diameter rings, whereas the other two used five rows containing 25 adjacent 25 cm diameter rings, respectively, all of which sampled the same underlying soil. It is notable that the variance of ln$i$ is approximately five times larger in the second Sisson and Wierenga study than in the third, but that it increases by a factor of only 1.5 between the second study and the first, suggesting that these last two measurements were correlated. This suggestion is consistent with their findings; they reported a correlation length of approximately 13 cm for the measurements made with the small rings.

A final point of note from this study is that the coefficient of infiltration rate variability for the small (6.2 × 6.2 m) study area is comparable to the rate observed in other studies on much larger fields, implying that much of the variability is present within a small structural unit. This conclusion was also drawn by Beckett and Webster (1971) in their summary of the spatial variation of solid-phase structural parameters and chemical nutrient concentrations.

In table 3, data summarize the variability of unsaturated hydraulic conductivity. In the studies reported, the hydraulic conductivity-water content function at each site was assumed to follow the exponential model

$$K(\theta) = K_o \exp[\beta(\theta - \theta_o)] \qquad [11]$$

(Libardi et al. 1980; Nielsen, Biggar, and Erh 1973), where K_o, β, and θ_o are fitted parameters not necessarily equal to the saturated values. Spatial variability in this function is manifest in the measured variability of the model parameters, K_o, β, and θ_o, the first two of which are given in table 3 for each study. The first four lines of table 3 are taken from the field study of Nielsen, Biggar, and Erh (1973). The first and third rows report the result of measuring $K(\theta)$ during redistribution at depths of 30 and 180 cm, respectively, by using the instantaneous profile method of Rose, Stern, and Drummand (1965). The second and fourth rows represent the result of a reanalysis of the same data set using the unit gradient method to measure $K(\theta)$ (Libardi et al. 1980). In both cases, the values of K_o and β were estimated from linear regression analysis of the log-transform of equation 11. This field measurement of $K(\theta)$ is notable because of the large variability of the K_o parameter (CV equal to 343 percent) as compared to that in the remaining studies and as compared to the variability of K_s measured on the same plots (CV equal to 243 percent, table 1). In contrast, the rest of the studies report a K_o variability comparable to the K_s variability summarized in table 1. Some of the large K_o variance found in the study by Nielsen, Biggar, and Erh (1973) could have resulted from poor agreement of the data with the model in equation 11.

Solute Transport Parameters

In discussing vertical solute transport through unsaturated soil, a distinction must be made between local-scale (i.e., field-plot) studies and field-scale studies (in which a description is sought of the area-averaged solute concentration at a given depth and time). In the former studies, the one-dimensional convection-dispersion equation

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} - V\frac{\partial C}{\partial z} \qquad [12]$$

is assumed to be valid locally, where C is the dissolved solute concentration, V is the

TABLE 3. FIELD STUDIES OF THE SPATIAL VARIABILITY OF UNSATURATED HYDRAULIC CONDUCTIVITY FUNCTION PARAMETERS FOR THE MODEL FUNCTION $K(\theta) = K_o \exp(\beta[\theta - \theta_o])$ (K_o in cm d^{-1})

Mean	$\sigma_{\ln K_o}$ or $\sigma_{\ln\beta}$	CV*	Texture	Area	Depth	n	Method of measurement	Reference
		%		*ha*	*cm*			
K_0 23	1.6	343	clay loam	150	30	20	instantaneous profile	Libardi et al. 1980
β 146	0.6	64						
K_0 4.6	1.4	235	clay loam	150	30	20	θ unit gradient method	Libardi et al. 1980
β 89	0.4	41						
K_0 34	1.4	263	clay loam	150	180	20	instantaneous profile	Libardi et al. 1980
β 62	0.7	73						
K_0 24	1.6	322	clay loam	150	180	20	θ unit gradient method	Libardi et al. 1980
β 56	0.4	47						
K_0 10	0.7	76	silt loam	150	0-160	611	unit gradient	Wagenet and Rao 1983
β 65	0.4	37						
K_0 4	0.4	46	loam	0.66	0-120	24	instantaneous profile on four plots	Simmons, Nielsen, and Biggar 1979
β 33	0.2	19						
K_0 10	0.7	79	sandy loam	0.50	30	99	θ unit gradient	Jones and Wagenet 1984
β 40	0.3	28						
K_0 10	0.6	62	sandy loam	0.50	120	81	θ unit gradient	Jones and Wagenet 1984
β 81	0.5	49						
K_0 11	0.8	91	sandy loam	0.50	0-120	641	θ unit gradient	Jones and Wagenet 1984
β 67	0.4	45						

*Coefficient of variation.

solute velocity, and D is the local longitudinal dispersion coefficient (Biggar and Nielsen 1976; Miller, Biggar, and Nielsen 1965; Warrick, Biggar, and Nielsen 1971). Application of equation 12 to different plots over a field (Biggar and Nielsen 1976), or to the effluent concentrations measured in a network of solution samplers (Jury, Stolzy, and Shouse 1982; Van De Pol, Wierenga, and Nielsen 1977) or soil cores (Jury, Elabd, and Collins 1983) after a pulse or front of solute has been added to the field, generates a set of V and D values that characterizes the lateral variability of one-dimensional solute transport. In field-scale solute transport, however, it is not evident that equation 12 is the correct physical law for describing area-averaged solute transport, either in unsaturated soil (Dagan and Bresler 1979; Jury 1983) or in groundwater (Gelhar, Gutjahr, and Naff 1979; Matheron and deMarsily 1980).

Field-scale solute transport in unsaturated soil has been conceptualized in two different ways: either as a set of parallel, noninteracting vertical soil columns each obeying equation 12 with its own D and V (Amoozegar-Fard, Nielsen, and Warrick 1982; Dagan and Bresler, 1979) or a transfer-function approach that regards the entire field as a single transport unit with entry and exit surfaces (Jury 1982; Jury, Sposito, and White 1986). In the latter approach, solute transport through the entire volume of soil between the soil surface and a given depth of observation is characterized by a PDF for solute travel times from the soil surface to the observation depth. In either approach, local differences in vertical velocity or travel time dominate the area-averaged longitudinal spreading of the solute as it moves downward, and the local dispersion coefficient values at a given site have essentially no influence (Amoozegar-Fard, Nielsen, and Warrick 1982; Dagan and Bresler 1979).

The few field-scale experiments that have replicated the downward movement of a surface-applied solute pulse or front of a chemical can give information about the distribution of vertical solute velocities (or travel times) to a reference depth, if solution samplers are used at the reference depth. Similarly, if deep soil cores are taken some time after the uniform surface application of a solute pulse, the distribution of depths reached by solute may be related to the variations in travel time or solute velocity. In virtually all cases where the solute velocity has been determined, the frequency distribution of parameters representing the velocity (or travel time) has been skewed, and can be represented better by a lognormal than a normal probability distribution. For a lognormal distribution of velocities or travel times to a reference depth, the variance of the logarithm of travel time and solute velocity should be equal. If the process of the solute moving downward is dominated by differences in local convection that are not diminished by lateral mixing, then the logarithm of the distribution of depths reached by the individual pulses also should have the same variance as the log travel time or log velocity. Thus, log variance parameters are suitable standardized variables for comparing the variability of downward solute transport.

The data presented in table 4 summarize the results of seven experiments in which enough measurements were taken to allow a determination of log variance parameters. The standard deviation parameters are relatively similar in all but two of the studies, the study of Biggar and Nielsen (1976), who measured solute movement under surface ponding conditions, and the study of Richter (1984), who measured solute in drainage from 36 small undisturbed lysimeters placed in two fields. Under ponded water application, the amount of water entering the soil at different locations is controlled by the saturated hydraulic conductivity, so the variance of solute travel time should be

TABLE 4. STANDARD DEVIATION AND COEFFICIENT OF VARIATION (CV) OF LOG-TRANSFORMED PARAMETERS INFERRED FROM VARIOUS FIELD EXPERIMENTS ON SOLUTE TRANSPORT

Parameter*	Soil	Area	Water application	Measurements	n	σ	CV†	Reference
							%	
1. $\ln V$	loamy sand (with pan)	4.6 × 6.1 m (4 plots)	ponding	solution samples to 240 cm	44	0.69	78†	Starr et al. 1978
2. $\ln V$	clay	8 × 8 m	tricklers	solution samples to 150 cm	24	0.57	61	Van De Pol, Wierenga, and Nielsen 1977
3. $\ln V$	loamy with clay	150 ha (20 plots)	ponding	solution samples to 180 cm	120	1.25	194†	Biggar and Nielsen 1976
4. $\ln I$	sandy	0.64 ha (14 sites)	rainfall	solution samples to 180 cm	70	0.57	61	Jury, Stolzy, and Shouse 1982
5. $\ln Zp$	sandy	0.64 ha (36 sites)	sprinkler	soil cores to 300 cm	36	0.35	36	Jury, Elabd, and Collins 1983
6. $\ln Zp$	loamy sand	3 × 3 m (8 plots)	sprinkler	soil cores to 80 cm	32	0.67	75	Wild and Babiker 1976
7. $\ln V$	sandy loam	30 × 30 m	sprinkler	undisturbed 190 cm deep lysimeters	36	1.26	204†	Richter 1984

*V = solute velocity; I = net applied water required to leach pulse to depth Z; Zp = depth of peak height of solute concentration in core.

†Coefficient of variation estimated from equation 13.

much higher. A high log variance, therefore, should not be considered a property of the field, but of the field and the water application method combined (Jury 1982). In Richter's lysimeter study, the soil was dry at the time of solute application and substantial early breakthrough was noted in all lysimeters even though the surface was not ponded.

If the travel time or velocity distributions are truly lognormal, the coefficient of variation of the distribution is given by the equation

$$CV = [\exp(\sigma^2) - 1]^{1/2} \quad [13]$$

(Aitcheson and Brown 1976). Equation 13 was used to calculate the coefficients of variation given in table 4 when the sample CV was not reported. However, as pointed out by Aitcheson and Brown (1976), equation 13 will produce a value that generally exceeds the sample CV of the data set from which it is calculated. The studies of Jury and coworkers (Jury, Elabd, and Collins 1983; Jury, Stolzy, and Shouse 1982) were conducted on the same field. The 1983 study involved soil coring and was conducted under relatively uniform sprinkler irrigation, whereas the earlier study was conducted under transient conditions, with moving solution intercepted by porous cup samplers during a winter of highly variable rainfall events. For that reason the coefficient of variation of the net applied water might be greater than the coefficient of variation of the drainage past the point of observation, which was not measured during the experiment.

INFLUENCE OF CORRELATION ON VARIANCE

The analytical methods used to estimate the mean values and variances presented in tables 1 to 4 did not take into account correlation between nearby measurements, but instead considered the measurements as spatially independent and uncorrelated. Disregard of the spatial structure of the medium may lead to biased estimates of the moments of the assumed PDF. The spatial structure of the medium should be identified and quantified prior to the estimation of the moments of the assumed PDF.

Generally, the procedure for selecting the most appropriate statistical model to describe the spatial structure of the medium is iterative (Kitanidis and Vomvoris 1983). According to Schweppe (1973), the use of measurements to develop a mathematical model involves four steps: (1) hypothesize the structure of the model; (2) estimate the parameters of the model; (3) test the validity of the model; and (4) if a test for validity fails, diagnose and correct the error, and then repeat steps 1, 2, and 3. As pointed out by Kitanidis and Vomvoris (1983), the problem of selecting the most appropriate model remains to some extent in the realm of engineering judgment. In practice, attention is usually restricted to certain classes of model, which are chosen for their practicality and versatility as well as for their performance in past applications. Within these classes the simplest or the lowest-order model in agreement with the available data usually is selected. Selection of the final model from the candidates in each class should be carried out by a systematic model discrimination procedure, such as selecting the model that minimizes the Akaike Information Criterion (AIC) (Hippel 1981).

Given the structure of a model, estimation of the parameters of the model from the

available data is a well-defined algorithmic problem. The most advanced parametric estimation methods are the maximum likelihood (ML), restricted maximum likelihood (RML), minimum-variance unbiased quadratic (MVUQ), and weighted least squares (WLS) procedures. Each of these estimation procedures was presented and evaluated by Kitanidis (1983) for the estimation of the parameters of a generalized polynomial covariance function using rainfall data, and by Kitanidis and Lane (1985) and Hoeksema and Kitanidis (1985) for the estimation of the parameters of an exponential covariance function and a linear drift function using data on the hydraulic properties of selected aquifers.

Once the parameters of a model have been obtained, the validity of the model should be tested. One possible validation test involves first transforming the residual measurement vector, $z(x) = u(x) - m(x)$, into a vector of uncorrelated residuals, $y(x)$, and then analyzing the statistical properties of $y(x)$. According to Kitanidis and Vomvoris (1983), when the covariance function C_Z of $z(x)$ is fitted, it can be decomposed into a product of a matrix C and its transpose C^T such that

$$CC^T = C_Z. \qquad [14]$$

This implies that

$$I = C^{-1}C_Z(C^{-1})^T, \qquad [15]$$

where I is the identity matrix. Then $y = y(x)$ is defined by the equation

$$y = C^{-1}z. \qquad [16]$$

From equations 14 and 15, the first two moments of y are

$$E[y] = C^{-1}\,E[z] = 0 \qquad [17]$$

and

$$E[yy^T] = C^{-1}E[zz^T]\,(C^{-1})^T = I. \qquad [18]$$

Thus the transformation in equation 15 creates a set of uncorrelated, unit-variance residuals, $y(x)$.

Under the assumption that $z(x)$ is distributed normally, $y(x)$ will be distributed normally. Under the hypothesis that the model is valid (i.e., that $z[x]$ is second-order stationary), the elements of $y(x)$ are independent standard normal variates. Thus, a statistical analysis of $y(x)$ will reveal the validity of the assumptions about $z(x)$. Kitanidis and Vomvoris (1983) discuss a procedure for determining whether a candidate set $y(x_i)$ of transformed data meets these criteria. According to their analysis, two conditions on $y(x_i)$ should be satisfied in order for the selected model to be valid: (1) normality, requiring that each $y(x_i)$ is a zero normal variate at the 95 percent significance level,

$$-2 < y(x_i) < 2 \qquad [19]$$

for all x_i; and (2) no correlation between variates, requiring that the product $y(x_i)\cdot y(x_j)$, where i is unequal to j, has zero mean and unit variance at the 95 percent significance level,

$$-2 < y(x_i)y(x_j) < 2 \qquad [20]$$

for all x_i and x_j where i is unequal to j. Finally, it is necessary to test whether the $y(x_i)$ values taken as a whole appear to conform to the hypothesized $y(x)$ unit normal distribution. The sum of squares (SSR) should follow a chi-square distribution with $m = n - p$ degrees of freedom,

$$\text{SSR} = \sum_{i=1}^{n} y(x_i)^2 \simeq \chi^2(n - p), \qquad [21]$$

where n is the sample size of $z(x)$ and p is the number of the parameters of the drift function fitted from the data. The normality of the $y(x_i)$ residuals may be tested also by the chi-square test or by the Kolmogorov-Smirnov (KS) test (Haan 1977).

Another possible model validation procedure is the cross-validation test coupled with the kriging technique, sometimes called "jackknifing" (Gambolati and Volpi 1979; Russo and Jury 1987a). The test is performed by suppressing each of the n observation points one at a time and producing an estimate u^* at that point using the remaining $n - 1$ data points and the kriging technique. The validity of the model is tested by analyzing the error $[u_i(x) - u^*(x)_i]$, where $i = 1, n$. Two conditions must satisfied for the selected model to be consistent theoretically: (1) there must be no systematic error (i.e., the mean ME of the reduced error RE vector, $\text{RE}(x_i) = [u^*(x_i) - u(x_i)]/\text{var}[u^*(x_i) - u(x_i)]^{1/2}$, is zero)

$$\text{ME} = \frac{1}{n} \sum_{i=1}^{n} \text{RE}(x_i) \simeq 0; \qquad [22]$$

and (2) the kriging variance, $\text{var}[u^*(x_i) - u(x_i)]$, is consistent with the corresponding error $[u^*(x_i) - u(x_i)]$. Thus, the mean square reduced error (MRE) must be unity:

$$\text{MRE} = \left[\frac{1}{n} \sum_{i=1}^{n} \text{RE}(x_i)^2\right]^{1/2} \simeq 1. \qquad [23]$$

In addition, an overall effective measure of the accuracy of the model is provided by the mean-square error (MSE),

$$\text{MSE} = \left\{\frac{1}{n} \sum_{i=1}^{n} [u^*(x_i) - u(x_i)]^2\right\}^{1/2}, \qquad [24]$$

which should be close to zero if the model is accurate. Under the assumption that the reduced error $\text{RE}(x_i)$ is distributed normally (Journel and Huijbregts 1978), a t-test may be performed to test the significance of ME = 0 and MRE = 1. A chi-square test or a KS test may be used to examine the assumption of a normal distribution for $\text{RE}(x_i)$.

APPLICATIONS TO FIELD DATA

Although the field investigations summarized in tables 1 to 4 characterize the frequency distribution properties of the sample sets measured in each study, few generalizations can be drawn by comparing different studies and their corresponding soil characteristics. To illustrate the additional information which may appear from a more comprehensive analysis, we will now concentrate on contrasting the characteristics of two different comprehensive field studies in which sufficient quantities of data were collected to permit an analysis of spatial structure.

The field studies of Nielsen, Biggar, and Erh (1973) and Russo and Bresler (1981) provide large sets of measurements of saturated hydraulic conductivity K_s for two field soils different in texture and in sampling area. The spatial distribution of $u = \ln K_s$ will be analyzed at each soil depth, first by viewing the $\ln K_s$ data as a realization of a two-dimensional, isotropic random function, and then by viewing the data as a realization of a three-dimensional, isotropic random function.

In the following analyses, two different models are used for the covariance function of $u = \ln K_s$: (1) an exponential model with a nugget effect,

$$C_u(h_{ij}) = C_n \delta_{ij} + C_o \exp(-h_{ij}/a); \qquad [25a]$$

and (2) a spherical model with a nugget effect,

$$C_u(h_{ij}) = C_n \delta_{ij} + C_o[1 - \frac{3}{2}(\frac{h_{ij}}{a}) + \frac{1}{2}(\frac{h_{ij}}{a})^3], \qquad [25b]$$

where $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq$ to j, and $C_n + C_o = C(0)$ is the variance of $\ln K_s$. For the deterministic component of $u = \ln K_s$, the general drift function model

$$m(x) = \sum_{l=0}^{K-1} f_l(x)\, \beta_l \qquad [26]$$

is used, where $f_l(x)$ for $l = 1$ to $l = K - 1$ are known polynomial functions of the spatial coordinates, $K = (p + 1)(p + 2)/2$ (p being the order of the polynomial), $f_o(x) = 1$, and β_l for $l = 0$ to $l = K - 1$ are the unknown drift parameters. Only two cases will be considered: zero order ($p = 0$ and $K = 1$) and first order ($p = 1$ and $K = 3$) polynomials. These are known as the constant and linear drift models, respectively. For simplicity, models fitted with an exponential covariance will be denoted *E,* and those fitted with a spherical covariance will be denoted *S.* A model fitted with a constant drift will be denoted *C,* and one fitted with a linear drift will be denoted *L.* Thus, for example, $S + L$ represents a spherical model with linear drift.

The method used to estimate the three parameters of the covariance function C_u (equation 25) is the restricted maximum likelihood (RML) estimation procedure described by Kitanidis and Lane (1985). When the data follow a joint Gaussian distribution, RML estimates are known to be asymptotically unbiased, minimally variant, consistent, and normally distributed, with a covariance matrix given by the inverse of the information matrix (Kendall and Stuart 1979). For data that are not normal, the procedure can be viewed as a fitting based on a weighted sum of squares of prediction errors (Kitanidis 1985). The covariance parameters are estimated by maximizing the likelihood of generalized increments, independently of the unknown mean,

which is filtered by a transformation method suggested by Kitanidis (1983). Given the fitted covariance of $u(x)$, C_u, the drift parameters, β_l (for $l = 0$ to $l = K - 1$) equation 26 may be estimated using a weighted least squares (WLS) procedure. If β is the vector of the drift parameters such that $E[u(x)] = X\beta$, then it can be estimated as

$$\hat{\beta} = (X^T C_u X)^{-1} X^T C_u{}^{-1} u, \qquad [27]$$

where X is the spatial coordinate matrix used in standard linear regression.

Validation of the fitted models is carried out by cross-validation tests and analyses of the uncorrelated residuals $y(x_i)$ (equation 16). These tests indicate whether a given model provides an adequate fit to a given set of $u = \ln K_s$. In addition, the Akaike Information Criterion

$$\text{AIC} = 2(\pounds + P) \qquad [28]$$

is used for model discrimination tests, where $\pounds$ is the negative log-likelihood for a fitted model and P is the total number of independently adjusted parameters within the model. The "best" model is the one that minimizes the AIC (Schweppe 1973).

Hamra Field

Using the air-entry permeameter method of Bouwer (1966), Russo and Bresler (1981) measured the saturated hydraulic conductivity K_s at 30 different sites and 4 different depths in a 0.8 ha plot of Hamra Red Mediterranean soil. According to their analysis, the values of K_s at a given soil depth appear to correlate spatially with an integral scale J (equation 9) ranging from 14 to 34 m.

Two-dimensional analyses

The data in table 5 summarize conventional statistical properties of $\ln K_s$ for each depth, estimated from a sample size of $n = 30$. As is customary in making such estimates, all calculations assume that the observations are spatially independent and uncorrelated. The KS normality test and the chi-square test were applied to evaluate

TABLE 5. ESTIMATES OF THE MEAN μ, VARIANCE σ^2, SKEW COEFFICIENT k_{sk}, AND COEFFICIENT OF KURTOSIS k_c, THE CALCULATED VALUES OF χ^2 (FOR 3 DEGREES OF FREEDOM), AND THE KS TEST STATISTIC D, FOR $u = \ln K_s$, AT FOUR DEPTHS (HAMRA FIELD)*

Depth	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	χ^2 (df)	D
m						
0.0-0.3	1.9403	0.3637	−1.932	7.084	5.6 (3)	0.1759
0.3-0.6	1.6793	0.1668	−0.572	2.253	10.4† (3)	0.1524
0.6-0.9	0.7193	0.8408	−0.928	3.084	7.6 (3)	0.1858
0.9-1.2	0.4613	1.2050	−0.851	2.859	15.2† (3)	0.1931

*Values of K_s are in μm per second.

†Null (normal) hypothesis is rejected at the 0.05 level of significance.

the hypothesis that the observations of $u = \ln K_s$ were drawn from a population with the normal distribution, $N(\mu,\sigma^2)$. In table 5, the calculated values of chi-square and the KS statistic $D = \max|\,\mathrm{CDF}_t - \mathrm{CDF}_o\,|$ are given, where CDF is the cumulative distribution function and the subscripts t and o denote theoretical and observed, respectively. Generally, the mean value of $\ln K_s$, the skew coefficient k_{sk}, and the kurtosis coefficient k_c decreased with depth, whereas the variance of $\ln K_s$ increased with depth. The chi-square test rejected the null (normal) hypothesis for the 0.3 to 0.6 m and the 0.9 to 1.2 m depths. However, the chi-square test is quite sensitive to the tails of the assumed distribution. The KS test does not reject the null hypothesis for any given depth.

Estimated parameters of the covariance functions modeling the field spatial structure are given in table 6. In general, for a given soil depth the fitted parameters depend on the model assumed. In a number of instances, the estimation procedures returned minimum or maximum permitted values of $\hat{C}_n$ or $\hat{a}$.

Results of validation tests of the models in table 6 are summarized in table 7. For a given soil depth, all four models generally appear to be consistent with the data as indicated by the values of MRE $\simeq$ 1 and ME $\simeq$ 0 (cross-validation test), the calculated chi-square statistic, and the sum of squares of residuals (SSR, the analysis of the uncorrelated residuals). Only for four cases—the $E + L$ model (0 to 0.3 m depth), and

TABLE 6. RESTRICTED MAXIMUM LIKELIHOOD ESTIMATES OF THE PARAMETERS OF THE COVARIANCE FUNCTION (EQUATION 25), ASSUMING SPHERICAL *S* OR EXPONENTIAL *E* COVARIANCE MODEL, AND CONSTANT *C* OR LINEAR *L* DRIFT MODEL FOR THE FOUR DIFFERENT SOIL DEPTHS (HAMRA FIELD)*

Depth	Model	$\hat{C}_n$		$\hat{C}_o$		$\hat{a}$	
m							
0.0-0.3	E + C	0†	(—)	0.5420	(0.477)	37.8	(38.9)
	S + C	0†	(—)	0.4588	(0.109)	54.7	(9.2)
	E + L	0†	(—)	0.2893	(0.179)	16.6	(15.1)
	S + L	0†	(—)	0.3385	(0.098)	43.4	(8.7)
0.3-0.6	E + C	0.0658	(0.059)	0.2234	(0.633)	150.0‡	(—)
	S + C	0.0619	(0.031)	0.1619	(0.150)	150.0‡	(—)
	E + L	0.0010	(0.374)	0.0877	(0.374)	0.10§	(—)
	S + L	0.0807	(0.036)	0.0098	(0.041)	42.7	(116.3)
0.6-0.9	E + C	0.0356	(0.105)	3.767	(1.93)	150.0‡	(—)
	S + C	0.0483	(0.105)	1.655	(1.33)	135.1	(81.6)
	E + L	0.0634	(0.113)	2.4000	(1.49)	150.0‡	(—)
	S + L	0.0099	(0.139)	0.4599	(0.233)	32.1	(10.4)
0.9-1.2	E + C	0†	(—)	1.956	(2.19)	50.9	(64.0)
	S + C	0†	(—)	1.233	(0.598)	48.6	(15.6)
	E + L	0†	(—)	3.249	(8.27)	84.8	(218)
	S + L	0†	(—)	0.962	(0.348)	43.7	(11.3)

*Values in parentheses are the standard error of estimation (SE). If the value is zero, there is no standard error.

†Here $\hat{C}_n = C_{nmin}$.

‡Here $\hat{a} = a_{max}$.

§Here $\hat{a} = a_{min}$. No SE value was calculated.

TABLE 7. ESTIMATES OF CROSS-VALIDATION TESTS*, ANALYSIS OF UNCORRELATED RESIDUALS†, AND THE AKAIKE INFORMATION CRITERION AIC FOR THE MODELS LISTED IN TABLE 6

		Cross validation			Analysis of uncorrelated residuals				
Depth	Model	ME	MSE	MRE	$\hat{\mu}$	$\hat{\sigma}^2$	χ^2 (*df*)	SSR	AIC
m									
0.0-0.3	E + C	0.0118	0.4332	1.029	−0.0942	0.991	7.6 (3)	29.01	227.6
	S + C	0.0399	0.4362	1.080	−0.1097	1.001	5.5 (3)	29.40	227.1
	E + L	0.0204	0.4519	1.177	−0.0881	0.993	26.1 (3)‡	27.00	226.1
	S + L	−0.0028	0.4354	1.039	−0.0823	0.968	7.6 (3)	26.31	224.6
0.3-0.6	E + C	0.0213	0.3499	1.209	0.0717	1.013	32.0 (3)‡	29.53	190.6
	S + C	0.0241	0.3453	1.241	0.0838	0.994	37.0 (3)‡	29.00	189.6
	E + L	−0.0149	0.3501	1.353	−0.1237	0.985	7.9 (3)‡	27.00	184.2
	S + L	0.0105	0.3695	1.268	0.1123	0.994	5.7 (3)	27.17	183.5
0.6-0.9	E + C	0.0197	0.5110	1.005	−0.0312	0.992	3.1 (3)	29.00	174.0
	S + C	0.0246	0.5081	0.984	0.2662	0.877	1.8 (3)	27.51	171.7
	E + L	0.0296	0.5182	1.032	0.1285	0.983	7.4 (3)	27.00	173.5
	S + L	−0.0192	0.5323	0.977	0.1796	0.968	6.5 (3)	27.01	168.7
0.9-1.2	E + C	0.0281	0.6299	0.917	0.1840	0.966	3.9 (3)	29.00	170.3
	S + C	−0.0083	0.5861	0.949	0.1608	1.013	2.2 (3)	30.10	167.2
	E + L	0.0436	0.6771	0.957	0.0357	0.988	3.4 (3)	26.72	174.3
	S + L	0.0322	0.6560	1.035	0.0610	0.998	0.8 (3)	27.06	212.4

*Mean error ME, mean square error MSE, mean reduced error MRE.

†Mean μ, variance σ^2, χ^2, sum of squares of residuals SSR.

‡Null (normal) hypothesis is rejected at the 0.05 level of significance.

the $E + C$, $S + C$, and $E + L$ models (0.3 to 0.6 m depth)—do the results of the chi-square test suggest rejection of the null hypothesis. Out of the models that pass the cross-validation test and the analysis of uncorrelated residuals (table 7) for a given soil depth, the AIC values are used to select the most appropriate model. The models that minimize the value of AIC (table 7) are the $S + L$ model for the three upper layers and the $S + C$ model for the 0.9 to 1.2 m soil depth.

The results in tables 6 and 7 demonstrate the necessity of using several estimation and validation procedures together with the AIC to discriminate between models that pass all other tests. Had the analysis used only the $E + C$ model and the cross-validation test, the final estimates of $\hat{a}$, $\hat{C}_o$, and $\hat{C}_n$ would have been significantly different from those actually chosen.

The parameters of the covariance and the drift functions describing the variability of $\ln K_s$ are summarized in table 8. The parameter $\hat{C}(0) = \hat{C}_n + \hat{C}_o$ is an estimate of the stochastic high-frequency variation (HFV) of $\ln K_s$. At the 0.3 to 0.6 m depth, about 90 percent of the HFV stems from variability at a scale smaller than the smallest separation distance in the horizontal plane of the field ($\simeq$ 2 m), which appears as "white noise" or a "nugget effect," C_n. There appears to be no nugget effect at the 0 to 0.3 m and the 0.9 to 1.2 m depths and a small nugget effect, $\hat{C}_n = 0.02\,\hat{C}(0)$, at the 0.6 to 0.9 m depth. On the other hand, the value of $\hat{C}_D$ given by the equation

$$\hat{C}_D = \sum_{i=1}^{n} [\hat{m}(x_i) - \hat{\mu}_u]^2/n \qquad [29]$$

is an estimate of the deterministic low-frequency variation (LFV) of $\ln K_s$. The magnitude of the LFV is about one-third the magnitude of the HFV at the 0 to 0.3 m

TABLE 8. ESTIMATES OF STOCHASTIC PARAMETERS* AND DRIFT PARAMETERS† OF THE MODELS DESCRIBING TWO-DIMENSIONAL SPATIAL VARIABILITY OF $\ln K_s$ (HAMRA FIELD)

	Stochastic parameters				Drift parameters			
Depth	$\hat{C}_n$	$\hat{C}_o$	$\hat{C}(0)$	$\hat{a}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{C}_D$
m								
0.0-0.3	0	0.3385	0.3385	43.4	2.8311	−0.0110	−0.0073	0.1223
						$R^{2‡} = 0.902$, SE[§] = 1.012		
0.3-0.6	0.0807	0.0098	0.0905	42.7	2.2867	−0.0101	1.37×10^{-5}	0.083
						$R^2 = 0.986$, SE = 1.039		
0.6-0.9	0.0099	0.4599	0.4698	32.1	2.3268	−0.0132	−0.0246	0.4165
						$R^2 = 0.8167$, SE = 0.801		
0.9-1.2	0	1.064	1.233	48.6	0.4153			

*Nugget variance C_n, covariance C_o, total stochastic variance $C(0)$, shape factor a; see equation 25.

†Polynomial coefficients β_0, β_1, β_2, and drift variance C_D.

‡R^2 refers to the regression of the drift function on the data.

§SE, standard error, refers to the regression of the drift function on the data.

depth, but is of the same order of magnitude as the HFV at the 0.3 to 0.6 m and 0.6 to 0.9 m soil depths. For a given soil depth, the total variance of $\hat{\sigma}^2$ of $\ln K_s$ is

$$\hat{\sigma}^2 = \hat{C}(0) + \hat{C}_D. \qquad [30]$$

The values of $\hat{\sigma}^2$ calculated for each of the soil depths are higher than the corresponding values calculated using conventional statistical methods (table 5) because the conventional analysis ignores spatial correlations and treats all observations as independent.

Three-dimensional analyses

The spatial variability of $\ln K_s$ has been analyzed in the horizontal plane by viewing it as a realization of a two-dimensional, isotropic stochastic function, considering different horizontal layers to be noninteracting. In this section, we will take $\ln K_s$ to be a realization of a three-dimensional, isotropic stochastic function, with all four depths used simultaneously to estimate the parameters of the covariance and the drift functions of $\ln K_s$.

The data presented in table 9 summarize the results of conventional statistical analysis, structural analysis, and model-validation tests (cross-validation test and analysis of the uncorrelated residuals). Both the KS test and the chi-square test reject the null (normal) hypothesis for $\ln K_s$. The structural analysis results (table 9b) show that, by considering the depth variations of $\ln K_s$, the estimated correlation scale of $\ln K_s$ becomes much smaller ($\hat{J} = 0.7$ m) than that estimated by considering only the lateral variations of $\ln K_s$ at a given soil depth ($\hat{J} = 18$ m). The characteristic vertical length of the sampling domain, however, is 1 m, as compared to a lateral length of 100 m. The cross-validation test results (table 9c) suggest that both the $E + C$ and the $S + C$ models are accurate and consistent with the data. The analysis of the uncorrelated residuals (table 9d) shows that the null hypothesis should be rejected for the $E + C$ model. Conversely, the KS normality test for the uncorrelated residuals, calculated by using the $S + C$ model, accepts the null hypothesis. Based on these results, we select the $S + C$ model, which minimizes AIC, as the model of choice for the three-dimensional spatial distribution of $\ln K_s$.

Thus far, a constant drift has been assumed. To examine the possibility that linear drift is present, 60 observation points (n) were selected randomly from the four depths, with the resultant three-dimensional random network permitting the use of a transformation matrix (Hoeksema and Kitanidis 1985) in the three-dimensional coordinate system. Table 10 contains a summary of the results of conventional statistical analysis, structural analysis, cross-validation tests, and analysis of the uncorrelated residuals for this case. In contrast to the results given in table 9 for all 120 points, both the KS and chi-square tests now accept the null hypothesis for $\ln K_s$. The structural analysis results show that the estimated parameters of the covariance function are quite sensitive to the model selected for describing the spatial variability of $\ln K_s$. For a given covariance model, linear drift considerably increases the estimated correlation scale relative to the case where a constant drift was assumed. Results of the cross-validation test and of the analysis of the uncorrelated residuals provide no reason to doubt the validity and adequacy of any of the four models.

According to Russo and Jury (1987a), values of MRE in excess of unity imply that

the estimated correlation scale a is larger than the theoretical one. Although the $E + L$ and $S + L$ models in table 1 have MRE > 1, the value of MRE is only required to lie within an acceptable range (e.g., within 95 percent confidence limits, $CL_{.95}$) of its theoretical value of unity. Since all of the values are within that range, the AIC criterion can be used to make the final model selection. Thus, the $S + L$ model, which had the lowest value for AIC, was chosen to describe the three-dimensional distribution of $\ln K_s$. The corresponding drift parameters are $\hat{\beta}_0 = 4.343$, $\hat{\beta}_1 = 1.045 \times 10^{-2}$, $\hat{\beta}_2 = -1.137 \times 10^{-2}$, and $\hat{\beta}_3 = -1.865$, where $R^2 = 0.85$ and SE = 1.02. A significant vertical drift component ($\hat{\beta}_3$) has been removed with this model. The values $\hat{C}_D = 0.5480$ and $\hat{C}(0) = 0.4805$ are the resultant estimates for the LFV and the HFV of $\ln K_s$ in the three-dimensional domain, and the total variance $\hat{\sigma}^2 = 1.0285$. According to this model, about 40 percent of the HFV is generated from variations occurring at a scale smaller than the smallest lag distance in the sampling network (0.3 m). The integral scale of $\ln K_s$ over the three-dimensional domain ($\hat{J} = \sqrt{(1/5)}\,a =$ 14.5 m) is about 78 percent of the average integral scale ($\hat{J} = 18.6$ m) derived from the two-dimensional distributions of $\ln K_s$ at each of the four different soil layers (table 8).

TABLE 9. ESTIMATES OF THE VARIABILITY OF $u = \ln K_s$ USING VARIOUS ANALYTICAL AND TESTING METHODS, USING JOINTLY ALL FOUR DEPTHS OF THE HAMRA FIELD ($n = 120$)*

a. Conventional statistical analysis†

$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D
1.2003	1.0331	−1.224	4.206	23.1 (3)‡	0.1340§

b. Structural analysis

Model	$\hat{C}_n$	$\hat{C}_o$	$\hat{a}$	AIC
E + C	0 (—)	1.095 (.178)	0.667 (.142)	848.1
S + C	0 (—)	1.043 (.167)	1.135 (.198)	841.9

c. Cross validation

Model	Mean error	Mean square error	Mean reduced error
E + C	−0.00342	0.6711	1.019
S + C	0.01047	0.6247	0.976

d. Analysis of the uncorrelated residuals

Model	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D	SSR
E + C	0.1872	0.8915	−0.997	3.995	18.3(3)‡	0.1131§	118.2
S + C	0.1274	0.9270	−1.07	4.159	15.8(3)‡	0.1081	119.0

*Symbols are defined in tables 5, 7, and 8. Values in parentheses are the standard error of estimation (SE). If the value is zero, there is no standard error.

†Values of K_s are in μm per second.

‡Null (normal) hypothesis is rejected at the 0.05 level of significance.

§Null (normal) hypothesis is rejected at the 0.10 level of significance.

TABLE 10. ESTIMATES OF THE VARIABILITY OF $u = \ln K_s$ USING VARIOUS ANALYTICAL AND TESTING METHODS, USING $n = 60$ RANDOMLY SELECTED VALUES OF u FROM THE FOUR DEPTHS OF THE HAMRA FIELD*

a. Conventional statistical analysis†

$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D
1.213	0.9364	−1.166	4.243	6.4 (3)	0.1180

b. Structural analysis

Model	$\hat{C}_n$	$\hat{C}_o$	$\hat{a}$	AIC
E + C	0 (—)	1.007 (.209)	0.655 (.278)	432.9
S + C	0 (—)	0.971 (.202)	1.298 (.353)	429.6
E + L	0.1864 (.051)	0.384 (.296)	19.74 (23.6)	388.2
S + L	0.1955 (.050)	0.285 (.140)	32.51 (13.7)	386.1

c. Cross validation

Model	Mean error	Mean square error	Mean reduced error
E + C	0.0115	0.892	1.108
S + C	0.0126	0.838	1.059
E + L	0.0444	0.613	1.236
S + L	0.0326	0.612	1.229

d. Analysis of the uncorrelated residuals

Model	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D	SSR
E + C	−0.0396	0.9963	−1.020	4.290	4.7 (3)	0.1076	58.86
S + C	−0.0453	0.9971	−1.108	4.417	5.1 (3)	0.1110	58.95
E + L	−0.0357	0.9984	−0.550	3.574	3.1 (3)	0.0736	55.98
S + L	−0.0356	0.9985	−0.560	3.599	1.0 (3)	0.0781	55.99

*Symbols are defined in tables 5, 7, and 8. Values in parentheses are the standard error of estimation (SE). If the value is zero, there is no standard error.

†Values of K_s are in μm per second.

Panoche Field

Nielsen, Biggar, and Erh (1973) measured the saturated hydraulic conductivity K_s under steady-state ponded infiltration conditions at 20 different sites and 6 different depths in a 150 ha clay loam field on an alluvial fan of the Panoche soil series. A ponded 6.2 × 6.2 m plot was used to measure the steady-state vertical flux, with the average value of duplicate tensiometers located in the center 2 m^2 of the plot at different soil depths used to estimate the vertical hydraulic gradient. Since at a given depth only 20 measurements of K_s are available, the spatial distribution of $u = \ln K_s$ was analyzed in the entire field domain by viewing $u = \ln K_s$ as a realization of a three-dimensional, isotropic random function.

In table 11, the data summarize the results of conventional statistical analysis,

structural analysis, the cross-validation test, and analysis of the uncorrelated residuals of $\ln K_s$. With the conventional statistical analysis, the chi-square test rejects the null hypothesis for $\ln K_s$, whereas the KS test accepts it. The structural analysis (using only the $E + C$ model) produced a relatively small correlation scale ($\hat{J} = \sqrt{2}\,\hat{a} = 9.65$ m) as compared with the field horizontal dimensions (1,000 m and 1,500 m in the N-S and the E-W directions, respectively) and a variance of $\hat{C}(0) = \hat{C}_n + \hat{C}_o = 2.212$, which is larger than the variance $\hat{\sigma}^2 = 1.932$ of the samples analyzed by conventional methods. About 5 percent of the variability of $\ln K_s$ occurs at a scale smaller than the smallest lag distance of the sampling network (0.3 m). The results of the cross-validation test and the analysis of the uncorrelated residuals suggest that the $E + C$ structural model is accurate and consistent with the data. Results of both the chi-square test and the KS test suggest that as opposed to the sample values of $\ln K_s$, the uncorrelated residuals of $\ln K_s$ are distributed normally (table 11d).

As in the case of the Hamra field, the most appropriate model for the three-dimensional spatial distribution of $\ln K_s$ was selected by randomly choosing 60 observation points (n) from the six depths. The data in table 12 summarize the results of both conventional and structural analysis, as well as model validation and discrimination for this case. Results of both the chi-square test and KS test accept the null hypothesis for $\ln K_s$ (table 12a). The structural analysis results demonstrate the sensitivity of the

TABLE 11. ESTIMATES OF THE VARIABILITY OF $u = \ln K_s$ USING VARIOUS ANALYTICAL AND TESTING METHODS, USING JOINTLY ALL SIX DEPTHS OF THE PANOCHE FIELD*

a. Conventional statistical analysis†

$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D
0.120	1.932	−0.571	3.202	9.9‡ (3)	0.0744

b. Structural analysis

Model	$\hat{C}_n$	$\hat{C}_o$	$\hat{a}$
E + C	0.1135 (0.047)	2.099 (.613)	6.823 (3.373)

c. Cross validation

Model	Mean error	Mean square error	Mean reduced error
E + C	−0.0245	0.537	1.113

d. Analysis of the uncorrelated residuals

Model	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D	SSR
E + C	−0.147	1.016	−0.167	3.663	4.5 (3)	0.0665	119.0

*Values in parentheses are the standard error of estimation (SE). Symbols are defined in tables 5, 7, and 8.

†Values of K_s are in μm per second.

‡Null (normal) hypothesis is rejected at the 0.05 level of significance.

estimated parameters of the covariance function to the type of model used to describe the spatial variability of $\ln K_s$. As was the case for the Hamra field, inclusion of linear drift in the structural model considerably increased the estimated correlation scale of the stochastic component, relative to the case where a constant drift was assumed. The chi-square test of the uncorrelated residuals rejected the null hypothesis only for the $E + L$ model. The KS test results and the values of SSR (equation 21), however, accepted the null hypothesis for the uncorrelated residuals associated with each of the four models. The cross-validation test suggested that only the $E + C$ model could be considered as accurate and consistent with the data. The $S + L$ model minimized the value of AIC. The value of MRE > 1 associated with the $S + L$ model suggests that the size of the estimated correlation scale might be reduced. Therefore, $\hat{a}$ was reduced until the value of MRE (equation 23) within the allowable range of variability. The value of

TABLE 12. ESTIMATES OF THE VARIABILITY OF $u = \ln K_s$ USING VARIOUS ANALYTICAL AND TESTING METHODS, USING $n = 60$ RANDOMLY SELECTED VALUES OF u FROM THE SIX DEPTHS OF THE PANOCHE FIELD*

a. Conventional statistical analysis†

$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D
0.193	2.140	−0.821	3.635	7.2 (3)	0.1108

b. Structural analysis

Model	$\hat{C}_n$	$\hat{C}_o$	$\hat{a}$	AIC
E + C	0.0414 (0.092)	2.012 (0.609)	4.033 (2.32)	441.7
S + C	0.0721 (0.078)	2.005 (0.621)	7.583 (3.88)	442.1
E + L	0.2291 (0.053)	1.772 (1.03)	249.5 (243.6)	431.5
S + L	0.2162 (0.080)	1.475 (0.55)	86.6 (242.3)	431.4

c. Cross validation

Model	Mean error	Mean square error	Mean reduced error
E + C	−0.0113	0.7284	1.220
S + C	−0.0145	0.7534	1.263‡
E + L	−0.1018	0.7282	1.427‡
S + L	−0.0814	0.6962	1.415‡

d. Analysis of the uncorrelated residuals

Model	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{k}_{sk}$	$\hat{k}_c$	$\chi^2(df)$	D	SSR
E + C	0.0528	0.9970	−0.435	2.388	2.9 (3)	0.073	58.99
S + C	0.0528	0.9972	−0.421	2.380	2.9 (3)	0.065	59.00
E + L	−0.0104	0.9992	−0.508	3.671	9.6§ (3)	0.126	56.00
S + L	0.0088	0.9999	−0.502	3.104	7.2 (3)	0.097	55.95

*Symbols are defined in tables 5, 7, and 8. Values in parentheses are the standard error of estimation (SE).

†Values of K_s are in μm per second.

‡Value of statistic is out of the allowable range of variability.

§Null (normal) hypothesis is rejected at the 0.05 level of significance.

$\hat{a} = 18.1$ m (and $\hat{J} = 8.10$ m) produced a mean and variance of the kriging-reduced error, RE(x_i), in the allowable range of variability: $-0.37 < \hat{\mu} < 0.27$ and $0.975 < \hat{\sigma}^2 < 2.09$. The resultant drift parameters are $\hat{\beta}_0 = 0.767$, $\hat{\beta}_1 = 1.27 \times 10^{-3}$, $\hat{\beta}_2 = -1.83 \times 10^{-4}$, and $\hat{\beta}_3 = 0.595$, where $R^2 = 0.65$ and SE = 0.79. The values of $\hat{C}(0) = 1.692$ and $\hat{C}_D = 0.3879$ are the resultant estimates for the HFV and the LFV of $\ln K_s$ in the three-dimensional domain. The resultant value of the AIC (433.3) is only slightly higher than the minimum value of AIC associated with the RML estimate of $\hat{a} = 86.6$ m. However, this value is still considerably smaller than the values of AIC associated with the constant drift models. The statistical analysis of the resultant uncorrelated residuals, ($\chi^2[3] = 3.6$, $D = 0.121$, and SSR = 48.3) suggested that the revised model (with $\hat{J} = 8.1$ and total variance of $\hat{\sigma}^2 = 2.08$) would produce normally distributed uncorrelated residuals at the 95 percent significance level ($-0.22 < \hat{\mu} < 0.27$, $0.54 < \hat{\sigma}^2 < 1.19$, and $\hat{\chi}^2_{.025} = 37 < \text{SSR} = 48.3 < \hat{\chi}^2_{2.975} = 78$), and thus may be considered to be accurate and consistent with the data.

One may conclude from the above analyses that $u = \ln K_s$ in the Hamra and Panoche fields may be regarded as realizations of three-dimensional, isotropic, stochastic functions whose distribution is $N[M(x), C_u(h)]$; that is, the functions are distributed normally with a spatially varying mean and stationary covariance. The RML procedure was used to estimate the parameters of $C_u(h)$, and a WLS procedure (equation 27) was used to estimate the parameters of $M(x)$. The most appropriate model to describe the spatial variability of $U(x,\omega)$ from among those meeting all other criteria was selected as the one that minimized the AIC. A spherical covariance function (equation 25a) with linear drift (equation 26, with $K = 3$) was selected for both fields by this procedure. The three-dimensional spatial correlation of $\ln K_s$ for the Hamra field is characterized by an integral scale of $\hat{J} = 14.5$ m. The LFV accounted for more than 50 percent of the total variability of $\ln K_s$ ($\hat{\sigma}^2 = 1.028$), and about 40 percent of the HFV stems from variations at a scale less than 0.3 m. The three-dimensional spatial correlation of $\ln K_s$ for the Panoche field is characterized by an integral scale of $\hat{J} = 8.1$ m. The LFV accounted for only 18 percent of the total variability of $\ln K_s$ ($\hat{\sigma}^2 = 2.08$), and about 13 percent of the HFV stems from variations at a scale less than 0.3 m.

The contrast between the two fields is interesting, showing that an integral scale of the same order may represent the spatial structure of $\ln K_s$ for soils of different textures and sampling areas. Despite the similar integral scale, however, the two fields have markedly different drift contributions, nugget variances, and total variances.

SUMMARY AND CONCLUSIONS

The preceding analysis has demonstrated how a statistical study of spatial structure may reveal field properties that are not apparent from a simple calculation of the statistical moments of the set of parameter measurements. Cases in point are the studies of Nielsen, Biggar, and Erh (1973) and Russo and Bresler (1981), both of which were best described by three-dimensional spherical covariance values with linear three-dimensional drift functions when judged by objective model validation criteria. The saturated hydraulic conductivity values had comparable integral scales for their underlying stochastic components (8.1 m and 14.5 m, respectively), but the Hamra

field of Russo and Bresler had a much larger deterministic drift component than did the Panoche field of Nielsen and coworkers. Furthermore, the residual stochastic component of the Hamra field possessed a large nugget variance (40 percent of total), whereas that of the Panoche field was small (13 percent of total).

This study has demonstrated the extreme sensitivity of the estimate of the integral scale parameter to the type of model used for the semivariogram, to the assumption of drift, and to the type of validation test used to confirm the estimation procedure. Failure to subject a data set to a comprehensive analysis such as that carried out in this study may result in the selection of integral scale values that have little meaning. It is notable that none of the early studies reviewed by Jury (1985) that reported integral scale values used any statistical validation procedures, relying instead only on simple semivariogram or autocorrelation model estimates of the field structural parameters. For this reason, the correlation noted by Jury (1985) between sample spacing and integral scale may be an artifact of estimation procedures that failed to analyze the real structure of the fields where such values were reported.

It is clear from this study that greater care must be taken in future studies of spatial structure, both to utilize more sophisticated estimation procedures and to analyze for drift components. In this study only the spatial distribution of the saturated hydraulic conductivity from two different fields was analyzed in detail. Stochastic analysis of water or solute transport in the unsaturated zone requires knowledge of the spatial distribution of soil hydraulic properties (soil hydraulic conductivity and water content–water potential functions) as well as the spatial distribution of the components of the dispersion tensor. This formidable task may be eased by the introduction of the scaling factor concept, which involves viewing a scaling factor as a stochastic variable so that the spatial variability of soil hydraulic properties may be described by a univariate parameter distribution instead of a multivariate parameter distribution. In part II of this paper, we analyze the spatial distribution of a scaling factor derived from the theory of microscopic similitude (Miller and Miller 1956).

ACKNOWLEDGMENT

The authors wish to thank the Electric Power Research Institute and the Southern California Edison Company for their partial financial support of this project.

LITERATURE CITED

AITCHESON, J., and J. A. C. BROWN
1976. *The lognormal distribution.* Cambridge, England: Cambridge University Press.
AMOOZEGAR-FARD, A., D. R. NIELSEN, and A. W. WARRICK
1982. Soil solute concentration distributions for spatially varying pore water velocities and apparent diffusion coefficients. Soil Sci. Soc. Am. J. 46:3-8.
BABALOLA, O.
1978. Spatial variability of soil water properties in tropical soils of Nigeria. Soil Sci. 126:269-79.
BAKR, A. A., L. W. GELHAR, A. L. GUTJAHR, and J. R. MacMILLAN
1978. Stochastic analysis of spatial variability in subsurface flows. 1. Comparison of one- and three-dimensional flows. Water Resour. Res. 14:263-71.
BEAR, J.
1972. *Dynamics of fluids in porous media.* New York: American Elsevier.
BECKETT, P. H. T., and R. WEBSTER
1971. Soil variability: A review. Soils Fert. 34:1-15.
BIGGAR, J. W., and D. R. NIELSEN
1976. Spatial variability of the leaching characteristics of a field soil. Water Resour. Res. 12:78-84.
BOUWER, H.
1966. Rapid field measurement of air entry value and hydraulic conductivity of soil as significant parameters in flow system analysis. Water Resour. Res. 2:729-38.
BRESLER, E., and G. DAGAN
1979. Solute dispersion in unsaturated heterogeneous soil at field scale. 2. Applications. Soil Sci. Soc. Am. J. 43:467-72.
1981. Convective and pore scale dispersive solute transport in unsaturated heterogeneous fields. Water Resour. Res. 17:1683-93.
1983. Unsaturated flow in spatially variable fields. 2. Application of water flow models to various fields. Water Resour. Res. 19:421-28.
BRESLER, E., and R. E. GREEN
1982. Soil parameters and sampling scheme for characterizing soil hydraulic properties of a watershed. Univ. Hawaii Tech. Rep. 148.
CAMERON, D. R.
1978. Variability of soil water retention curves and predicted hydraulic conductivities. Soil Sci. 126:364-71.
CASSEL, D. K.
1983. Spatial and temporal variability of soil physical properties following tillage of Norfolk loamy sand. Soil Sci. Soc. Am. J. 47:196-201.
CRESSIE, N.
1984. Towards resistant geostatistics. In *Geostatistics for natural resources characterization, I. Proceedings of the NATO Advanced Study Institute,* eds. Verly et al. NATO ASI Series, Series C, Mathematical and Physical Sciences. Dordrecht, The Netherlands: D. Reidel Publishing.
CRESSIE, N., and D. M. HAWKINS
1980. Robust estimation of the variogram, I. J. Int. Assoc. Math. Geol. 12:115-25.
DAGAN, G.
1982. Stochastic modelling of groundwater flow by unconditional and conditional probabilities. 2. The solute transport. Water Resour. Res. 18:835-48.
1984. Solute transport in heterogeneous porous formations. J. Fluid Mech. 145:151-77.
DAGAN, G., and E. BRESLER
1979. Solute dispersion in unsaturated heterogeneous soil at field scale. I. Theory. Soil Sci. Soc. Am. J. 43:461-67.
1983. Unsaturated flow in spatially variable fields. 1. Derivation of models of infiltration and redistribution. Water Resour. Res. 19:413-20.
DELFINER, P.
1976. Linear estimation of non-stationary spatial phenomena. In *Advanced geostatistics in the mining industry,* eds. M. Guarasco et al., 49-68. Dordrecht, The Netherlands: D. Reidel Publishing.
DELHOMME, J. P.
1978. Kriging in the hydrosciences. Adv. Water Resour. 1:251-66.
1979. Spatial variability and uncertainty in ground-water flow parameters: A geostatistical approach. Water Resour. Res. 15:269-80.

DOOB, J. L.
1953. *Stochastic processes.* New York: John Wiley.
DOWD, P. A.
1984. The variogram and kriging: Robust and resistant estimators, In *Geostatistics for natural resources characterization, I. Proceedings of the NATO Advanced Study Institute,* eds. Verly et al. NATO ASI Series, Series C, Mathematical and Physical Sciences. Dordrecht, The Netherlands: D. Reidel Publishing.
DUFFY, C., P. J. WIERENGA, and R. A. KSELIK
1981. Variations in infiltration rate based on soil survey information and field measurements. N. M. Agric. Exp. Stn. Bull. 60.
FREEZE, R. A.
1975. A stochastic conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. Water Resour. Res. 11:725-41.
GAMBOLATI, G., and G. VOLPI
1979. Groundwater contour mapping in Venice by stochastic interpolators. 1. Theory. Water Resour. Res. 15:281-90.
GELHAR, L. W.
1986. Stochastic subsurface hydrology from theory to applications. Water Resour. Res. 22:135S-45S.
GELHAR, L. W., and C. AXNESS
1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. Water Resour. Res. 19:161-80
GELHAR, L. W., A. L. GUTJAHR, and R. L. NAFF
1979. Stochastic analysis of macrodispersion in a stratified aquifer. Water Resour. Res. 15:1387-97.
GREMINGER, P. J., Y. Y. SUD, and D. R. NIELSEN
1985. Spatial variability of field-measured soil-water characteristics. Soil Sci. Soc. Am. J. 49:1075-82.
GUMAA, G. A.
1978. *Spatial variability of in situ available water.* Ph.D. dissertation. Univ. of Arizona, Tucson.
HAAN, C. T.
1977. *Statistical methods in hydrology.* Ames: The Iowa State University Press.
HALD, A.
1952. *Statistical theory with engineering applications.* New York: Wiley.
HIPPEL, K. W.
1981. Geophysical model discrimination using the Akaike Information Criterion. IEEE Trans. Autom. Control AC-26(2):358-78.
HOEKSEMA, R. J., and P. K. KITANIDIS
1985. Analysis of the spatial structure of properties of selected aquifers. Water Resour. Res. 21:563-72.
JONES, A. J., and R. J. WAGENET
1984. *In situ* estimation of hydraulic conductivity using simplified methods. Water Resour. Res. 20:1620-26.
JOURNEL, A. G., and Ch. J. HUIJBREGTS
1978. *Mining geostatistics.* London: Academic Press.
JURY, W. A.
1982. Simulation of solute transport with a transfer function model. Water Resour. Res. 18:363-68.
1983. Chemical transport modeling: Current research and unresolved problems. In *Chemical mobility and reactivity in soil systems.* Soil Sci. Soc. Am. Special Publ. II.
1985. Spatial variability of soil physical parameters insolute migration: A critical literature review. EPRI Topical Rep. EA4228. Palo Alto: Electric Power Research Institute.
JURY, W. A., H. ELABD, and T. J. COLLINS
1983. Field scale movement of adsorbing and nonadsorbing chemicals applied to the soil surface. In *Proc. Symp. Characterization and Monitoring of the Vadose Zone,* 203-22. Worthington, Ohio: National Water Well Assn.
JURY, W. A., G. SPOSITO, and R. E. WHITE
1986. A transfer function model of solute transport through soil. 1. Fundamental concepts. Water Resour. Res. 22:243-47.
JURY, W. A., L. H. STOLZY, and P. SHOUSE
1982. A field test of the transfer function model for predicting solute transport. Water Resour. Res. 18:369-75.
KEISLING, T. C., J. M. DAVIDSON, D. L. WEEKS, and R. D. MORRISON
1977. Precision with which selected soil physical parameters can be estimated. Soil Sci. 124:241-48.
KENDALL, M., and A. STUART
1979. *The advanced theory of statistics,* vol. 2. London: Charles Griffin & Co.

KITANIDIS, P. K.
1983. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. Water Resour. Res. 19:909-21.
1985. Minimum-variance unbiased quadratic estimation of covariances of regionalized variables. J. Int. Assoc. Math. Geol. 17:195-208.
KITANIDIS, P. K., and R. W. LANE
1985. Maximum likelihood parameter estimation of hydrologic spatial processes by the Gauss-Newton method. J. Hydrol. 79:53-71.
KITANIDIS, P. K., and E. G. VOMVORIS
1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. Water Resour. Res. 19:677-90.
LIBARDI, P. L., K. REICHARDT, D. R NIELSEN, and J. W. BIGGAR
1980. Simple field methods for estimating soil hydraulic conductivity. Soil Sci. Soc. Am. J. 44:3-6.
LUMLEY, J. L., and A. PANOFSKY
1964. *The structure of atmospheric turbulence.* New York: John Wiley & Sons.
MATHERON, G.
1971. The theory of regionalized variables and its applications. Cah. Cent. Morphol. Math. 5:211.
MATHERON, G., and G. deMARSILY
1980. Is transport in porous media always diffusive? A counterexample. Water Resour. Res. 16:901-7.
MILLER, E. E., and R. D. MILLER
1956. Physical theory for capillary flow phenomena. J. Appl. Phys. 27:324-32.
MILLER, R. J., J. W. BIGGAR, and D. R. NIELSEN
1965. Chloride displacement in Panoche clay loam in relation to water movementand distribution. Water Resour. Res. 1:63-73.
NIELSEN, D. R., J. W. BIGGAR, and K. T. ERH
1973. Spatial variability of field-measured soil-water properties. Hilgardia 42(7):215-59.
PANCHEV, S.
1971. *Random functions and turbulence.* Oxford: Pergamon Press.
RICHTER, G.
1984. *Microlysimeter and field study of water and chemical movement through soil.* M. S. thesis. University of California, Riverside.
ROGOWSKI, A. S.
1972. Watershed physics: Soil variability criteria. Water Resour. Res. 8:1015-23.
ROSE, C. W., W. R. STERN, and J. E. DRUMMAND
1965. Determination of hydraulic conductivity as a function of depth and water content for soil *in situ.* Aust. J. Soil Res. 3:1-9.
RUSSO, D.
1984a. Design of an optimal sampling network for estimating the variogram. Soil Sci. Soc. Am. J. 48:708-16.
1984b. A geostatistical approach to solute transport in heterogeneous fields and its applications to salinity management. Water Resour. Res. 20:1260-70.
RUSSO, D., and E. BRESLER
1981. Soil hydraulic properties as stochastic processes. I. An analysis of field spatial variability. Soil Sci. Soc. Am. J. 45:682-87.
1982. Soil hydraulic properties as stochastic processes. II. Errors of estimates in a heterogeneous field. Soil Sci. Soc. Am. J. 46:20-6.
RUSSO, D., and W. A. JURY
1987a. A theoretical study of the estimation of the correlation scale in spatially variable fields. I. Stationary fields. Water Resour. Res. (in press).
1987b. A theoretical study of the estimation of the correlation scale in spatially variable fields. II. Nonstationary fields. Water Resour. Res. (in press).
SCHWEPPE, F. C.
1973. *Uncertain dynamic systems.* Englewood Cliffs: Prentice-Hall.
SHARMA, M. L., G. A. GANDER, and C. G. HUNT
1980. Spatial variability of infiltration in a watershed. J. Hydrol. 45:101-22.
SIMMONS, C. S.
1982. A stochastic convective transport representation of dispersion in one-dimensional porous media systems. Water Resour. Res.18:1193-1214.
SIMMONS, C. S., D. R. NIELSEN, and J. W. BIGGAR
1979. Scaling of field-measured soil-water properties. Hilgardia 47(4):77-174.
SISSON, J. B., and P. J. WIERENGA
1981. Spatial variability of steady state infiltration rates as a stochastic process. Soil Sci. Soc. Am. J. 45:699-704.

SMITH, L., and F. W. SCHWARTZ
1980. Mass transport. 1. Stochastic analysis of macrodispersion. Water Resour. Res. 16:303-13.
SPOSITO, G., W. A. JURY, and V. K. GUPTA
1986. Fundamental problems in the stochastic convection-dispersion model of solute transport in aquifers and field soils. Water Resour. Res. 22:77-88.
STARR, J. L., H. C. DE ROO, C. R. FRINK, and J. Y. PARLANGE
1978. Leaching characteristics of layered field soil. Soil Sci. Soc. Am. J. 42:386-91.
VAN DE POL, R. M., P. J. WIERENGA, and D. R. NIELSEN
1977. Solute movement in a field soil. Soil Sci. Soc. Am. J. 41:10-13.
VIEIRA, S. R., D. R. NIELSEN, and J. W. BIGGAR
1981. Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 1040-48.
WAGENET, R. J.
1985. Measurement and interpretation of spatially variable leaching processes. In *Soil spatial variability,* eds. D. R. Nielsen and J. Bouma. Proc. of Workshop of ISSS and Soil Sci. Soc. Am., Wageningen, Netherlands: Pudoc.
WAGENET, R. J., and B. K. RAO
1983. Description of nitrogen movement in the presence of spatially variable soil hydraulic properties. Agric. Water Manage. 6:227-43.
WARRICK, A. W., J. W. BIGGAR, and D. R. NIELSEN
1971. Simultaneous solute and water transfer for an unsaturated soil. Water Resour. Res. 7:1216-25.
WARRICK, A. W., and D. R. NIELSEN
1980. Spatial variability of soil physical properties in the field. In *Applications of soil physics,* ed. D. Hillel, 319-44. New York: Academic Press.
WILD, A., and I. A. BABIKER
1976. The aysmmetric leaching pattern of NO_3 and Cl in a loamy sand under field conditions. J. Soil Sci. 27:460-66.
WILLARDSON, L. S., and R. L. HURST
1965. Sample size estimates in permeability studies. J. Irrig. Drain. Div. ASCE 91:1-9.

Continued from inside front cover.
component than the Panoche field of Nielsen, Biggar, and Erh (1973). The stochastic component of $\ln K_s$ in the Bet-Dagan field possessed a large nugget variance (40 percent of total) and was characterized by an integral scale of $J = 14.5$ m, as compared with $J = 8.1$ m and a small nugget variance (13 percent of total) in the Panoche field.

II. Scaling Models of Water Transport

In this paper, we examine the possibility of introducing a single stochastic scaling factor α, derived from macroscopic Miller similitude, to describe the spatial variability of soil hydraulic properties. Most of the information available allowed only a conventional statistical analysis of the scaling factors derived from different soil properties. The field studies of Nielsen, Biggar, and Erh (1973) and Russo and Bresler (1981) were suitable also for more detailed structural analyses. Results of these analyses suggested that the spatial structure of the α-set derived from the hydraulic conductivity function $K(\theta)$ is different from that of the α-set derived from the water retentivity function $h(\theta)$, reflecting the different spatial structures of the $K(\theta)$ and the $h(\theta)$ functions. Consequently, the statistical relationship between the uncorrelated residuals of the two α-sets was rather weak. For the Hamra field of Russo and Bresler (1981), the use of relative hydraulic properties to estimate the scaling factor sets considerably improved the correlation between the α-sets, which had essentially the same spatial structure but slightly different variances.

In this study, where the soil hydraulic properties are assumed to be described by the model of Brooks and Corey (1964), analytical expressions for the variances of the two different α-sets indicated that (1) both α-sets are dependent on the range of water saturation that is used to estimate them, (2) the correlation between the two sets will improve in media with a wide pore-size distribution, and (3) the two sets will be identical if and only if the relative hydraulic conductivity function $K_r(h_r)$ is described by a deterministic function, $K_r(h_r) = h_r^{-2}$. This result suggested that, in general, a second scaling factor for K_r is required for media that are not characterized by this single deterministic relationship.

A more general $K_r(h_r)$ relation, defined by $K_r = h_r^{-\eta}$, was introduced using η as a second stochastic variable. In this representation, the α scaling factor for K_r is defined by $K_r/K_r^* = \alpha^\eta$ instead of α^2 as in macroscopic Miller similitude. For the Hamra field, the resultant new α-set was identical to the α-set derived from the relative retentivity function. For the Panoche field, using the values of η to estimate the scaling factor from the relative hydraulic conductivity function considerably improved the correlation and the similarity between the two α-sets, but did not render them identical. The results of our analysis suggest that, for transient water flow, describing the spatial variability of $K(\theta)$ and $h(\theta)$ requires at least three stochastic variates: K_s, α, and η.

The Journal HILGARDIA is published irregularly. Number of pages and number of issues vary per annually numbered volume. Address: Agriculture and Natural Resources Publications, University of California, Berkeley, CA 94720.