WATER RESOURCES RESEARCH, VOL. 20, NO. 6, PAGES 682–690, JUNE 1984

A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils

B. J. COSBY, G. M. HORNBERGER, R. B. CLAPP, AND T. R. GINN

Department of Environmental Sciences, University of Virginia

Stochastic modeling of soil water fluxes in the absence of measured hydraulic parameters requires a knowledge of the expected distribution of the hydraulic parameters in different soil types. Predictive relationships describing the hydraulic parameter distributions must be developed based on the common descriptors of the physical properties of soils (e.g., texture, structure, particle size distribution). Covariation among the hydraulic parameters within these relationships must be identified. Data for 1448 soil samples were examined in an evaluation of the usefulness of qualitative descriptors as predictors of soil hydraulic behavior. Analysis of variance and multiple linear regression techniques were used to derive quantitative expressions for the moments of the hydraulic parameters as functions of the particle size distributions (percent sand, silt, and clay content) of soils. Discriminant analysis suggests that the covariation of the hydraulic parameters can be used to construct a classification scheme based on the hydraulic behavior of soils that is analogous to the textural classification scheme based on the sand, silt, and clay content of soils.

INTRODUCTION

Application of the classical theory of soil water movement requires knowledge of the relationships among matric potential, moisture content, and hydraulic conductivity. The physical attributes of the soil giving rise to these interrelationships are understood in a qualitative sense [e.g., *Childs*, 1969]. A comprehensive theory to allow derivation of the relationships from fundamental properties of the medium (e.g., grain size distribution) is, however, not yet fully developed, although recent work suggests that certain aspects of the hydraulics may be amenable to a theoretical treatment [*Nakano*, 1976; *Arya and Paris*, 1981]. In most cases, curves of matric potential versus moisture content (the moisture characteristic) and of hydraulic conductivity versus either matric potential or moisture content must be determined for a given soil by direct measurement. Statistical analyses can be used to identify what soil properties are important in describing the observed variation in these curves, thereby providing information of practical value as well as suggesting how theoretical exploration might proceed.

One approach that has been used to define the moisture characteristic is the construction of regression equations to predict the moisture content at specified values of matric potential using properties such as bulk density, percent sand, and other measured properties such as organic matter content [*Ghosh*, 1980; *Gupta and Larson*, 1979; *Rawls and Brakensiek*, 1982]. Results from these studies indicate that reasonable predictions can be made when the necessary data are available.

An alternate approach that has proven useful when data on grain size distribution are not available is to parameterize the moisture characteristic and then to investigate parameter variability with respect to soil physical properties. *Brakensiek et al.* [1981] and *McCuen et al.* [1981] examined the Brooks-Corey and Green-Ampt parameters using data from *Holtan et al.* [1968] and *Rawls et al.* [1976]. *Brakensiek et al.* [1981] examined the distribution of these parameters across textural classes defined on the U.S. Department of Agriculture (USDA) soil triangle. Various transforms of the data were applied to arrive at normal distributions of the parameters across all textural classes. The means and standard deviations of each parameter within a textural class were reported for each transform. Correlations among the parameters within a textural class were also given. No attempt was made to determine whether a regular pattern of variation in the parameters occurred across textural classes, and no explanation was offered for the observed correlations within classes. *McCuen et al.* [1981] established that the Brooks-Corey and Green-Ampt parameters differ significantly across textural classes. They also reported means, standard deviations, and simple correlations for the parameters within each textural class. The parameter statistics were presented overlain on the USDA textural triangle. The authors concluded that while there were trends obvious in the variations of the parameters over the triangle, there were numerous "irrational" results and concluded that a clear answer could not be obtained regarding the systematic individual variation of the parameters. They then examined the collective variation of the parameters using multivariate analysis of variance followed by a discriminant analysis. These results indicated that a weighted combination of the parameter values (i.e., a discriminant score) showed a more rational variation over the textural triangle. Again no attempt was made to relate the observed variation of means and standard deviations to the the physical properties of the textural classes. They emphasized that while the tabulated statistics of the individual parameters for each class provided a useful approximation to the hydraulic behavior of the soils, these statistics ignored important interrelationships in the parameters. *Clapp and Hornberger* [1978] also analyzed a portion of the data. They noted that the slope of the moisture characteristic curve was correlated with the clay fraction of the textural class.

The present paper provides an extension of the work described above. First, we wanted to determine if there was significant variation of the soil moisture parameters with physical properties of the soil other than texture. Second, we wanted to quantify, if possible, any observed relationships between the statistical properties of the parameters and the physical properties of the soils. Third, we wanted to extend the investigation begun by *McCuen et al.* [1981] into the interrelationships among the parameters.

Paper number 4W0237.
0043-1397/84/004W-0237$05.00

TABLE 1. List of Sample and Site Descriptors and the Classes for Each Descriptor

Descriptor	Classes
Texture	sand (14), sandy loam (124), loamy sand (30), loam (103), silty loam (394), sandy clay loam (104), silty clay loam (325), clay loam (147), sandy clay (16), silty clay (43), light clay (148)
Horizon	A (488), B (795), C (165)
Moist consistency	very friable (248), friable (643), firm (390), very firm (74), unclassified (93)
Structural size	very fine (66), fine (520), medium (560), coarse (129), unclassified (173)
Structural form	platy (50), prismatic (113), blocky (176), subangular blocky (621), granular (337), crumbly (13), massive (98), unclassified (40)
Roots	abundant (220), common (345), few (314), none (269), unclassified (300)
Topography (local slope)	0–2% (402), 2–7% (735), 7–14% (220), 14–25% (58), 25–55% (16), unclassified (17)
Drainage	very poor (27), poor (65), somewhat poor (161), moderate (337), well (794), somewhat excessive (27), excessive (33), unclassified (4)
Land use	long-term pasture (628), long-term cultivated (629), long-term forest (124), long-term idle (67)

The number in parentheses is the number of samples in each classification. Texture, land use, and horizon were available for all samples. Other descriptors were not always available for each sample. Unclassified samples were not included in statistical analyses using that descriptor.

DATA AND METHODS

The data are from *Holtan et al.* [1968] and *Rawls et al.* [1976]. The soil samples used to generate these data were taken from 35 localities in 23 states in the United States. In each testing area, several sampling sites were chosen and all horizons were subsampled. For each subsample the following hydraulic data are available: (1) moisture retention on a weight-weight basis determined at 0.1, 0.3, 0.6, 3.0, and 15.0 bars using ceramic plate and membrane techniques, (2) bulk density measured by displacement of the sample dried to 0.3 bar tension, (3) saturated hydraulic conductivity determined (usually in duplicate) in the laboratory using a 1-inch slice of a fist-sized fragment trimmed to roughly cylindrical shape. Details of the methods used are given by *Holtan et al.* [1968] and *Rawls et al.* [1976]. The weight-weight moisture retention data were converted to volume-volume measures (Θ) for each matric potential (Ψ), and the saturated water content (Θ_s) was determined for each sample using the bulk density and assuming a specific gravity of 2.65 for all solids. All matric potentials were converted to centimeters of water.

We chose to use what we consider to be a minimal set of parameters to describe the hydraulic properties. Two of these, the saturated hydraulic conductivity K_s and the saturated moisture content Θ_s are measured quantities in the data set. The other two (Ψ_s and b) are derived by fitting a power function,

$$\Psi = \Psi_s(\Theta/\Theta_s)^b \tag{1}$$

to the moisture retention data. The two derived parameters are thus Ψ_s, the "saturation" matric potential, and b, the slope of the retention curve (on a logarithmic graph). The b exponent and K_s can be used to estimate the entire hydraulic conductivity–moisture content curve [*Campbell*, 1974].

Forms other than (1) have been used to represent the moisture characteristic. The most widely used of these is the one from *Brooks and Corey* [1964]. That equation requires estimation of an additional parameter, the residual saturation. *Brakensiek* [1979] points out that the formulation which includes the residual saturation "generally gives a better fit to the moisture retention data." We argue that the limited number of measurements taken for each sample (five values of Θ and Ψ) and the large amount of variability in the available data suggest that a simpler representation of the hydraulic properties is desirable for our purposes. Also, some studies indicate that the power function form is entirely adequate [e.g., *Ghosh*, 1980]. Thus we use (1).

For each sample, values of log Ψ_s and b were determined by taking the logarithm of both sides of (1) and performing a linear regression. A preliminary analysis of the results indicated that Θ_s, log Ψ_s, and b were approximately normally distributed over all of the samples. No further transformations of these data were undertaken before the statistical analysis. If duplicate measurements of K_s were available for a sample, a geometric mean of the values was used. The K_s values were log transformed before the statistical analyses since they were highly skewed.

The combined data sets contained 1873 soil samples. Only those samples were used for which moisture characteristics and saturated conductivities were both available. Additionally, samples texturally classified as rock fragments or identified as R horizon were deleted. This resulted in 1448 samples for analysis. No further a priori selection of the data was attempted.

For each of the 1448 samples, descriptions of the physical properties of the soil and characteristics of the sampling site are available [*Holtan et al.*, 1968; *Rawls et al.*, 1976]. Each descriptor consists of several classes; every set of hydraulic parameters was assigned to one class of each descriptor based on the information in the data set. The descriptors and their classes are summarized in Table 1. Once the hydraulic parameters and descriptor classifications had been determined for all samples, the analysis proceeded in four stages.

First, a one-way analysis of variance was performed for each descriptor to determine if the hydraulic parameters varied significantly over the classes of that descriptor. That is, we wanted to determine if patterns existed in the individual

TABLE 2. Values of Percent Silt, Sand, and Clay Content Used for Each Textural Class in the Regression Analyses

Class	Percent Silt	Percent Sand	Percent Clay
Sand	5	92	3
Loamy sand	12	82	6
Sandy loam	32	58	10
Loam	39	43	18
Silty loam	70	17	13
Sandy clay loam	15	58	27
Clay loam	34	32	34
Silty clay loam	56	10	34
Sandy clay	6	52	42
Silty clay	47	6	47
Clay	20	22	58

The percentages were obtained from midpoint values of each textural class using the USDA textural triangle.

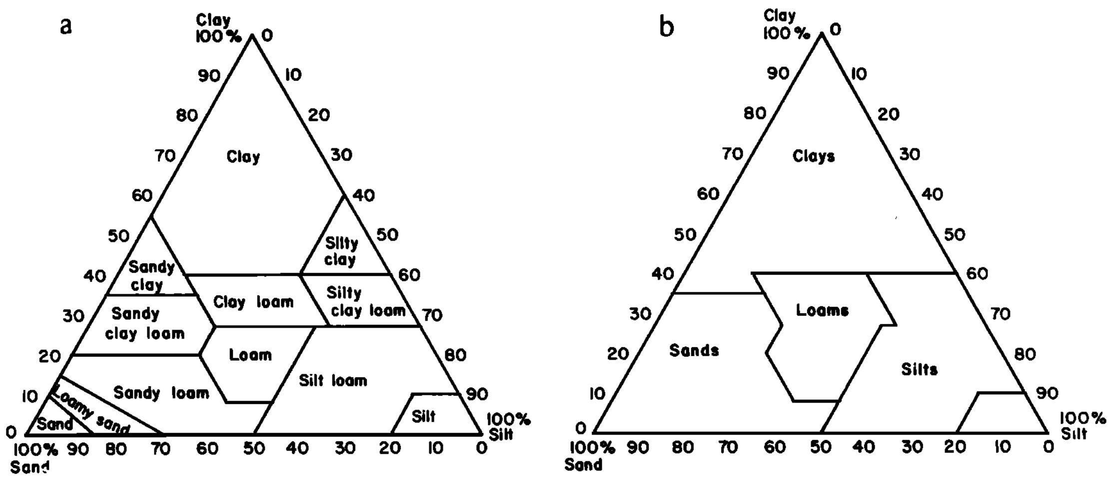

Fig. 1. (*a*) The USDA soil texture triangle. (*b*) Reclassification of the texture categories into four broad regions for the two and four group discriminant analyses.

hydraulic parameters that could be described by what was known of the physical properties of the sample or site. Since all of the descriptors are qualitative rather than numerical, a nonordinal technique (analysis of variance) was applied at this stage. That is, correlations or regressions between parameter values and descriptor classes cannot be calculated since it is not possible to associate sensible numerical values with many of the descriptor classifications. Instead, the analysis of variance provides a means to determine whether or not the parameter distributions change from one class of a descriptor to another. Once a descriptor that is associated with variation in a parameter has been identified, further attempts can be made to quantify the relationship (see below).

Second, a two-way analysis of variance was performed to determine if there was overlapping information about the hydraulic parameters contained in the soil or site descriptors identified in the first analysis. For instance, one would expect that texture and structure would be closely related, and if a given parameter varied significantly over textural classes, it would be expected to vary also over structural classes. In an analysis of variance (either one or two way) the fraction of the total variance in a parameter that is attributable to membership in classes of some descriptor can be estimated. If two descriptors, each identified by a one-way analysis of variance as being important, are included simultaneously in a two-way analysis and if the proportion of parameter variance attributable to class memberships in the two-way analysis is essentially the same as that attributable to class membership of either descriptor alone, then the information contained in the two descriptors is redundant. In such a case, either descriptor will suffice to describe all that can be known of the parameter variation. We decided (see results) that a single descriptor, texture, can account for most if not all of the discernible patterns in the individual parameters.

These results led to the third stage of the analysis, an attempt to quantify the pattern over the textural classes to provide a predictive relationship for the hydraulic parameters. Although *Holtan et al.* [1968] did assign a textural class to each sample, no actual particle size distribution data were available. We adopted the approach of *Clapp and Hornberger* [1978] and assigned values of percent silt, sand, and clay to each textural class based on the midpoint values of each textural class on the *U.S. Department of Agriculture* [1951, p. 209] textural triangle. These percentages are given for each textural class in Table 2; the triangle is reproduced in Figure 1*a* (Figure 1*b* will be referred to in the results section). Using these percentages for each textural class, a multiple linear regression analysis was performed using the average value of each parameter (or log-transformed parameter) within a given textural class as the dependent variable and the 11 sets of size fraction data in Table 2 as the independent variables. A second multiple linear regression analysis was performed using the standard deviations of each parameter within a class as the dependent variable and the percentages in Table 2 as the independent variables. Knowing not only the mean but also the variance of a parameter within a textural class as a function of

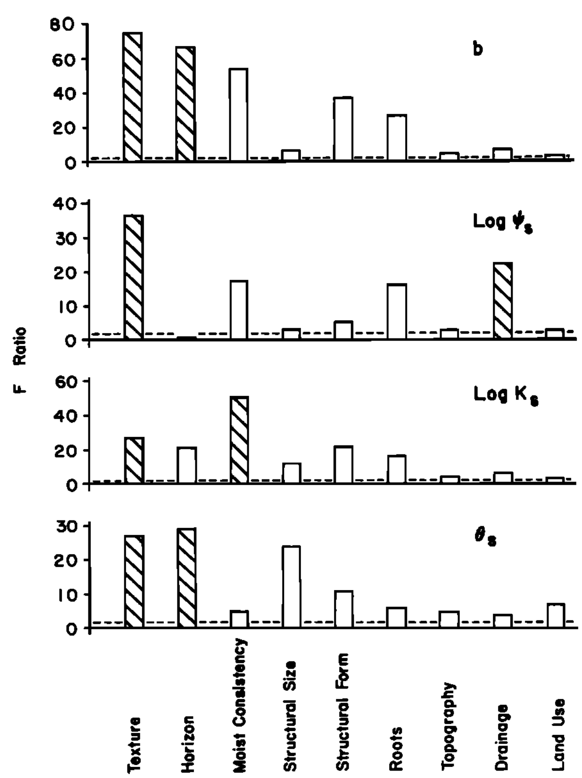

Fig. 2. Values of the F ratios from the one-way analyses of variance. The dashed line represents a significant result ($p = 0.10$).

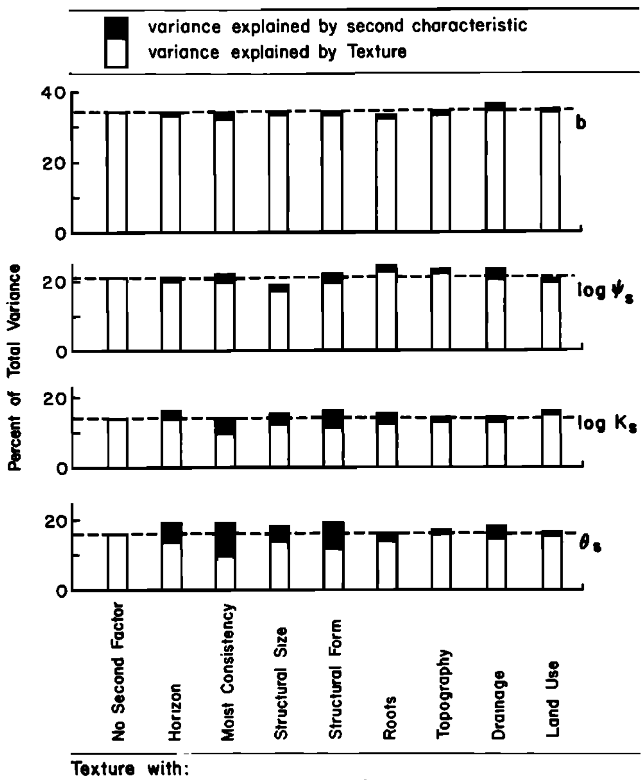

Fig. 3. Percent of variance explained by each factor in the two-way analyses of variance. The dashed line represents the percent of variance explained by texture alone.

sand, silt, or clay content has obvious advantages in interpreting or modeling soil water variability.

The analyses in the first three stages were concerned only with how each hydraulic parameter individually may depend on the physical properties of the soil. There exists the obvious possibility that the parameters covary and that some linear (or nonlinear) combination of the parameters are related to the physical properties of the soil. The fourth stage of the analysis attempted to address this question of covariability of the hydraulic parameters. A discriminant analysis using the hydraulic parameters was performed with the discriminant groups defined by textural classes. The resulting discriminant functions (i.e., weighted linear combinations of the hydraulic parameters) can be considered to define a new space (by defining new coordinate axes) that contains not only the information derived from the univariate analyses but also the important interactions of the original parameters. The discriminant space, as shown below, displays a striking resemblence to the textural space defined by the silt, sand, clay triangle, further reinforcing the results of the univariate analysis.

We should point out that we are here interested in an exploratory statistical analysis of the data and not in a conventional hypothesis-testing analysis. "Data-dredging" procedures [*Selvin and Stuart*, 1966] are often used in the examination of data sets not collected as part of an experiment to test a specific hypothesis. Such procedures may be useful for suggesting hypotheses (to be tested using independently collected data), but the strict interpretation of statistical tests may be inappropriate. Thus we are concerned only with exposing "robust" (i.e., well-defined) relationships in the data; precise measures of significance are not of concern. Violations of the assumptions of analysis of variance are therefore not crucial in this work, particularly since these violations (e.g., heteroscedasticity) will only result in reduced efficiency of estimation and not in bias [*Kendall and Stuart*, 1968].

The statistical analyses were performed using the *Statistical Package for the Social Sciences* [*Nie et al.*, 1975]. The routines were implemented on a CDC CYBER 730-2 at the University of Virginia. A discussion of all methods used can be found in the works by *Cooley and Lohnes* [1971], *Kendall and Stuart* [1968], *Nie et al.* [1975], and *Tatsuoka* [1971].

Results

One-Way Analysis of Variance

The results of the one-way analysis of variance (ANOVA) are presented graphically in Figure 2. The height of each bar represents the value of the F ratio derived from the one-way ANOVA for each parameter for each descriptor. Since F is the ratio of the parameter variance between groups to the parameter variance within groups, a large value of F indicates a significant change in the parameter distribution from class to class of the descriptor. The number of degrees of freedom for each test can be calculated from the data given in Table 1. The

TABLE 3. Means and Standard Deviations for the Four Hydraulic Parameters in Each Textural Class

		b		log Ψ_s		log K_s		Θ_s	
Class	n	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Sandy loam	124	4.74	1.40	1.15	0.73	−0.13	0.67	43.4	8.8
Sand	14	2.79	1.38	0.84	0.56	0.82	0.39	33.9	7.3
Loamy sand	30	4.26	1.95	0.56	0.73	0.30	0.51	42.1	7.2
Loam	103	5.25	1.66	1.55	0.66	−0.32	0.63	43.9	7.4
Silty loam	394	5.33	1.72	1.88	0.38	−0.40	0.55	47.6	5.4
Sandy clay loam	104	6.77	3.39	1.13	1.04	−0.20	0.54	40.4	4.8
Clay loam	147	8.17	3.74	1.42	0.72	−0.46	0.59	46.5	5.4
Silty clay loam	325	8.72	4.33	1.79	0.58	−0.54	0.61	46.4	4.6
Sandy clay	16	10.73	1.54	0.99	0.56	0.01	0.33	40.6	3.2
Silty clay	43	10.39	4.27	1.51	0.84	−0.72	0.69	46.8	6.2
Light clay	148	11.55	3.93	1.67	0.59	−0.86	0.62	46.8	3.5
All classes	1448	7.22	3.86	1.59	0.70	−0.42	0.64	45.7	6.1

Parameters: b is the slope of log Ψ versus log (Θ/Θ_s) regression, Ψ in centimeters H_2O; log Ψ_s is the intercept of log Ψ versus log (Θ/Θ_s) regression, Ψ in centimeters H_2O; log K_s is the log of the saturated hydraulic conductivity in inches per hour; Θ_s is the saturated water content in percent (volume/volume).

dashed line in Figure 2 represents a significance level of $p = 0.10$ for each result. Examination of the figure shows that at the $p = 0.10$ level, all parameters show significant variation on all descriptors (with the exception of log Ψ_s analyzed with horizon). Finding a significant result for all descriptors is not surprising given the large size of the data set. We are, however, interested only in robust relationships (i.e., very large values of F) between the hydraulic parameters and the soil descriptors. By this criterion, the two shaded bars for each parameter represent the two most important descriptors for that parameter. In all cases, texture is one of the two most important.

Two-Way Analysis of Variance

This led us to examine the degree of overlap in information between the descriptors using a two-way analysis of variance with texture always one of the two categories. The results of the two-way ANOVA are presented in Figure 3. The height of each bar represents the percent of the total variance in each parameter attributable to membership in the classes of texture and each other descriptor. The dashed line is equal to the percent of the total variance attributable to membership in the classes of texture alone (from the preceding one-way analysis of variance). For each bar representing texture with some other descriptor, the unshaded portion represents the percent of the parameter variance attributable to texture when that descriptor is entered first in the analysis, and the shaded portion represents the amount of residual variance attributable to the additional descriptor. Several things are apparent from Figure 3. The proportion of the total variance in a parameter that can be attributed to a combination of texture and some other descriptor is essentially the same in all cases as that attributable to texture alone. This implies that the information about parameter variability that each descriptor other than texture contains (see Figure 2) is redundant information. For the available set of data, texture alone should suffice to describe all that can be known, in practical terms, of the parameter variability.

In the few cases where the second descriptor explains a sizable proportion of the variance, the results must be interpreted cautiously. For example, consider the analysis of Θ_s with texture and moist consistency. In this case, the explained variance is roughly equally divided between texture and moist consistency. However, the total variance explained is only slightly greater than the variance explained by texture alone.

TABLE 4. Results of Multiple Linear Regression Analyses on the Means and Standard Deviations of the Parameters

Parameter	Intercept	Variable	Slope	R^2	ΔR^2	p	n
Mean b	3.10	% clay	0.157	0.966		0.001	11
		% sand	−0.003	0.966	0	0.769	
Mean log Ψ_s	1.54	% sand	−0.0095	0.809		0.001	11
		% silt	0.0063	0.850	0.041	0.180	
Mean log K_s	−0.60	% sand	0.0126	0.839		0.001	11
		% clay	−0.0064	0.872	0.033	0.193	
Mean Θ_s	50.5	% sand	−0.142	0.771		0.001	11
		% clay	−0.037	0.785	0.014	0.484	
S.D. b	0.92	% clay	0.0492	0.524		0.012	11
		% silt	0.0144	0.584	0.060	0.314	
S.D. log Ψ_s	0.72	% silt	−0.0026	0.096		0.355	11
		% clay	0.0012	0.111	0.015	0.716	
S.D. log K_s	0.43	% silt	0.0032	0.369		0.047	11
		% clay	0.0011	0.403	0.034	0.519	
S.D. Θ_s	8.23	% clay	−0.0805	0.567		0.007	11
		% sand	−0.0070	0.574	0.007	0.721	

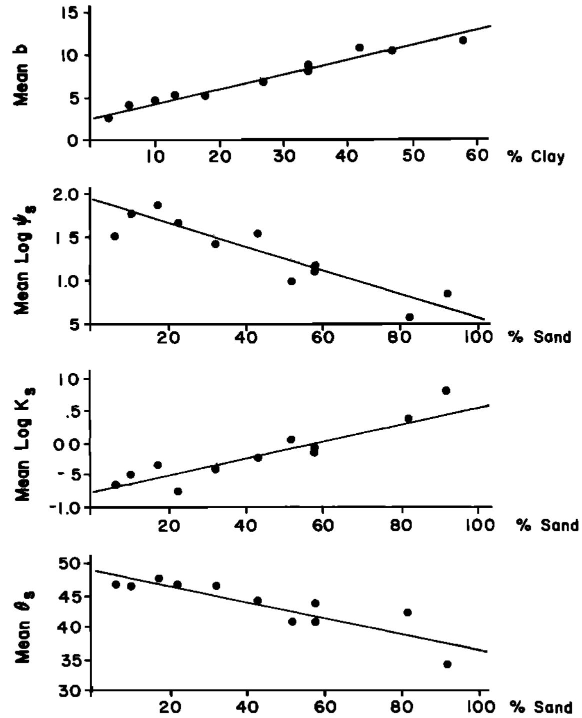

Fig. 4. Plots of the mean values of the hydraulic parameters for each textural class versus the most important variable (percent sand, silt, or clay) determined from the multiple linear regression analysis. The solid line is the univariate regression line.

The total explainable variance is fixed by the data; the apportioning of that variance to each descriptor when the information in each descriptor is redundant will be determined by the design of the analysis and may vary as the design varies. Also note that while all 1448 samples were assigned to a textural class, some of the samples were not classified on the other physical descriptors (e.g., moist consistency, see Table 1). This resulted in different degrees of freedom for each two-way ANOVA and is responsible for the different apportioning of the variance and the fact that in some cases the total explained variance in an analysis containing texture with a second descriptor is slightly less than the variance explained by texture alone. Put another way, several of the two-way ANOVA's in Figure 3 were performed on a subsample of the total data set and thus cannot be expected to apportion the variance identically to an analysis performed on the entire data set. Nonetheless, by the previous criterion of robustness it is apparent from Figure 3 that the additional information from a second descriptor beyond that provided by texture alone is marginal.

Multiple Linear Regression Analysis

As a first step in examining the dependence of the parameters on textural class, the means and standard deviations of each parameter for each textural class were calculated. The values are given in Table 3. Multiple linear regression (MLR) analysis was performed on the means and standard deviations of each parameter using the percent sand, silt, and clay values for each textural class as the independent variables. Note that there are really only two independent variables for each class since the sum of percent sand plus silt plus clay must equal 100. The MLR analysis was designed to pick the most important variable (in the sense of most parameter variance ex-

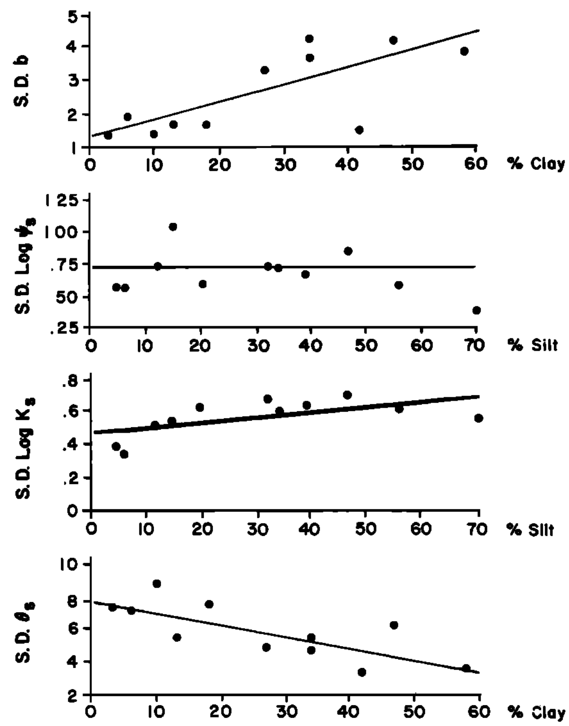

Fig. 5. Plots of the standard deviations of the hydraulic parameters within each textural class versus the most important variable (percent sand, silt, or clay) determined from the multiple linear regression analysis. The solid line is the univariate regression line.

plained by the regression) from sand, silt, or clay and, having corrected for the linear relationship in that variable, select the second most important variable from the two remaining. This procedure was also applied to the raw data for each class to check the regressions on the parameter means of each class. The slopes and intercepts for regressions using mean values (11 classes) were essentially the same as those using the raw data (1448 samples). No such check could be performed for the standard deviations of each class.

Table 4 summarizes the results of the MLR procedure. The table gives the intercept of the multiple regression, the most important (top) and second most important (bottom) variable, the regression slope for each variable, the R^2 (proportion of sum of squares) value for the regression when only the first variable is included (top) and when both are included (bottom), the increase in R^2 as a result of adding the second variable, and the significance of each variable in the regression. As can be seen, for all cases except the standard deviation of log Ψ_s, there was only one significant variable in the regression ($p = 0.10$). The standard deviation of log Ψ_s had no significant relationship to percent sand, silt, or clay. The ANOVA results for log Ψ_s depended solely on the fact that the mean of log Ψ_s varied over textural classes. For the other three parameters, however, both the means and standard deviations of the parameters varied as a function of soil textural class. This result has not been reported in other analyses of these data, and its importance will be discussed below. Figures 4 and 5 show the class means and standard deviations of each parameter plotted against the most important variable, percent sand, silt, or clay, as determined by the MLR procedure. The solid lines included in Figures 4 and 5 are the univariate (not multivariate) regression lines for each parameter on the most important variable. The slopes, intercepts, r^2 values, and significances of these univariate regressions are presented in Table 5. The univariate regression equations are very similar (but not identical) to the multivariate results in Table 4. The similarity derives from the fact that the second variable in each of the multivariate regressions is not very important (in the sense that the increase in R^2 due to the second variable is small relative to the overall R^2 value). By the previous criterion of robustness, a univariate regression of each parameter should be sufficient to describe most of the variability in hydraulic parameters over textural classes. The univariate results in Table 5 represent predictive relationships for the hydraulic parameters based on knowledge of the physical properties of soils. Using the multivariate regressions as predictive relationships results in only a marginal increase in information.

To assess the power of these regression relationships to explain the variability in each parameter, we returned to the original data set. For each individual soil sample the measured or calculated values of the four hydraulic parameters were normalized by subtracting the mean and dividing by the standard deviation of each class using the reported textural class and the univariate or multivariate regression equations. The resulting normalized parameter values should be independent of textural class if the univariate dependences shown in Figures 4 and 5 or the multivariate dependences of Table 4 are removed from the data. Another one-way analysis of variance was performed on the normalized parameters. The results are shown in Figure 6 which is a plot of the ANOVA F ratios for each parameter before (a) and after normalization using (b) the univariate regression equations and (c) the multivariate regression equations. All F ratios are significant at a level of $p = 0.10$. In all cases, the normalized parameters are more uniformly distributed over the textural classes (smaller F ratios indicate less dependence of the parameter on textural class). Additionally, the figure indicates that for all parameters, with the possible exception of log K_s, using the univariate regression relationships to describe parameter variation over texture is just as good as using the multivariate relationships.

While the regression equations apparently account for much of the variability of the hydraulic parameters over different soils, the F ratios of the normalized parameters are still significant (albeit much reduced). We can speculate that the remaining variability of the parameters could be reduced if the exact particle size distribution for each sample were known rather than the approximate values based on the midpoint of the given textural class. However, our original intention was to develop a predictive relationship based on qualitative soil descriptors. There are several alternate explanations of the remaining variability. In particular, it may be that soil properties not only affect each parameter individually but also affect the covariation of the parameters in a manner not completely

TABLE 5. Results of the Univariate Regressions of the Hydraulic Parameters on Percent Sand, Silt, or Clay

Parameter	Variable	Slope	Intercept	r^2	Significant at $p = 0.10$
Mean b	% clay	0.159	2.91	0.966	yes
Mean log Ψ_s	% sand	−0.0131	1.88	0.809	yes
Mean log K_s	% sand	0.0153	−0.884	0.839	yes
Mean Θ_s	% sand	−0.126	48.9	0.771	yes
S.D. b	% clay	0.0500	1.34	0.524	yes
S.D. log Ψ_s	···	···	···	···	no
S.D. log K_s	% silt	0.00321	0.459	0.369	yes
S.D. Θ_s	% clay	−0.0730	7.73	0.567	yes

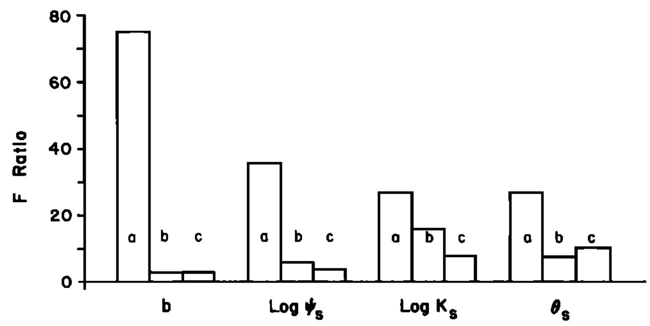

Fig. 6. F ratios from the one-way analysis of variance for each parameter, before (*a*) and after normalization of the parameter values using (*b*) the univariate expressions of Table 5 and (*c*) the multivariate expressions of Table 4. All F ratios are significant ($p = 0.10$).

described by the individual regression relationships. To examine this possibility, we "inverted" the problem. That is, rather than attempting to find some numerical property of texture that can predict the parameter values, we attempted to find some property of the parameter values that can predict the textural class of the sample. This property, for instance, a sum or product of the four parameter values, would depend on percent sand, silt, and clay just as textural class depends on those variables. Proceeding from the simplest case, we decided to examine a weighted linear sum of the hydraulic parameters. The weight for each parameter can be chosen to maximize the variability of the sum over the textural classes using a classical discriminant analysis procedure.

Discriminant Analyses

The soil textural triangle was divided into the four regions indicated in Figure 1*b*. The textural classes within each region were lumped into four broad categories, sand, silt, clay, and loam, for the first discriminant analyses; the lower right corner of the triangle is ignored since there were no samples labeled with the textural class "silt." Three initial discriminant analyses were performed: sand versus all others, silt versus all others, and clay versus all others. The analyses each contained two discriminant categories, and therefore only one discriminant function was derived in each analysis. Analysis A in Table 6 gives the correlation between the discriminant scores and the parameter values for all samples in the data set (the remainder of Table 6 is discussed below). These correlations maybe thought of as the importance of each parameter in the particular weighted linear combination of the parameters that best differentiates between a given particle size class and all others.

The discriminant analysis was next performed on all four broad discriminant categories simultaneously. This design allowed for the calculation of three discriminant functions; however, only two were significant at the $p = 0.10$ level (significance determined by Wilks' lambda). The two sets of discriminant score/parameter value correlations from the four category analysis are presented in analysis B in Table 6. The two functions accounted for 99.8% of the explainable variance in the data.

A final detailed discriminant analysis was performed using all 11 textural classes as the discriminant categories. Since there were four discriminant variables, four functions were derived. All four functions were significant at the $p = 0.10$ level (Wilks lambda); however, the first two functions accounted for 97.2% of the explained variance. Therefore only the first two functions are considered. The discriminant score/parameter value correlations are presented in analysis C in Table 6.

For the two category analyses (analysis A), the highest correlations for b and log K_s occur on function DCL, which discriminates clays from all the rest. This can be interpreted as meaning that soils rich in clay can best be discriminated from soils poor in clay by the slope of the moisture characteristic and the saturated hydraulic conductivity of a soil sample. The relationship of b and clay content was already known from the univariate regression analysis. The saturated matric potential Ψ_s and porosity Θ_s of the soil are important in differentiating soils rich in sands and silts from other soils but are relatively unimportant in discriminating clay-rich soils. The important fact is that all hydraulic parameters have significant weights on all functions (except possibly for log Ψ_s and Θ_s on DCL), and therefore we must conclude that the hydraulic uniqueness of the three basic soil types, sands, silts, and clays, arises from combinations of the hydraulic parameters and that they cannot be characterized by any single hydraulic parameter.

Returning now to the relationship of the hydraulic parameters to textural class, we can attempt to relate the broad (four category) and detailed (11 category) discriminant results to the distinguishing characteristics of the three basic soil particle size classes. Notice that the two important functions for both the four and 11 category analyses are very similar. The pattern of parameter variation over the textural classes is robust and appears at both coarse and fine scales. To interpret the discriminant functions from the four and 11 group cases, we calculated the correlations between the discriminant scores based on textural groupings (D4A, D4B, D11A, and D11B) and the discriminant scores from the analyses based on particle sizes (DCL, DSN and DSL). The results are presented in Table 7. For both the four and 11 group analyses, the second discriminant functions (D4B and D11B) are highly correlated with the function which best discriminates silts from sands and

TABLE 6. Correlation Coefficients (r) Between Canonical Discriminant Function Scores and the Hydraulic Parameter Values

Analysis	Analysis Design	Discriminant Categories	Discriminant Function	Hydraulic Parameters			
				b	log Ψ_s	log K_s	Θ_s
A	2 categories	clay vs. all others	DCL	−0.92	0.01	0.44	−0.08
		sand vs. all others	DSN	0.31	0.51	−0.36	0.43
		silt vs. all others	DSL	0.19	−0.76	0.14	−0.48
B	4 categories	sand, silt, clay, loam	D4A	0.41	0.45	−0.37	0.37
			D4B	0.85	−0.57	−0.22	−0.30
C	11 categories	the 11 textural classes	D11A	0.51	0.33	−0.38	0.30
			D11B	0.79	−0.55	−0.03	−0.43

TABLE 7. Correlation Coefficients (r) Between Discriminant Function Scores

Two-Category Discriminant Function	Four-Category Discriminant Function		Eleven-Category Discriminant Function	
	D4A	D4B	D11A	D11B
DCL	−0.79	−0.61	−0.80	−0.59
DSN	0.96	−0.27	0.96	−0.29
DSL	0.10	−0.99	0.08	−0.99

clays (the coefficients are −0.99 for both correlations). This suggests that we might interpret the second discriminant function as a silt axis. The first function for each group might likewise be interpreted as a clay-sand axis. The two discriminant functions can be used to define a planar parameter space similar to the sand, silt, clay planar space defined by the USDA triangle. Figure 7 shows a plot of the two discriminant function scores for the four group case. The functions were evaluated for each group using the mean of the parameter values within each group. Superimposed on the discriminant space is the modified USDA triangle from Figure 1*b*. The similarity between the classification scheme based on weighted combinations of the hydraulic parameters and the classification scheme based on particle size distributions of the soils is striking.

A similar plot of the discriminant scores based on the 11 group analysis is presented in Figure 8. A distorted version of the textural triangle (Figure 1*a*) is superimposed on the discriminant space. Again, the similarity between the two spaces is striking. While the relative areas of the textural classes have changed, the neighbor-to-neighbor relationship is identical in the two spaces. The only textural class which falls outside its expected region is the sandy clay class. It should be noted, however, that this class was represented by only 16 samples in the total sample population of 1448.

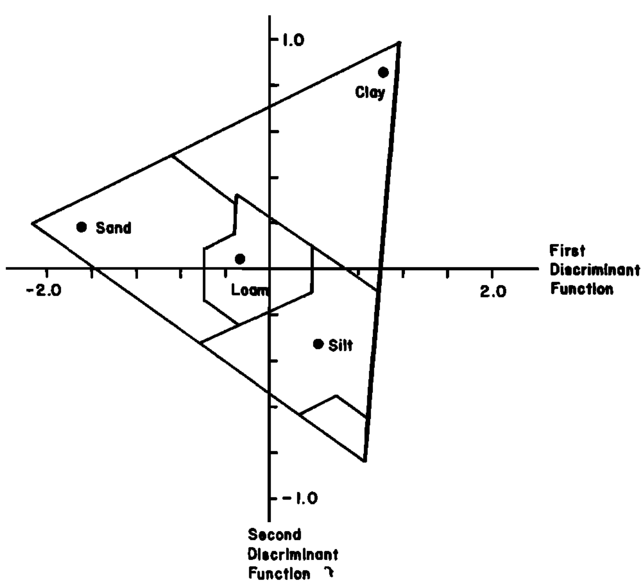

Fig. 7. Plot of the two significant discriminant function scores for the four group analysis. The modified USDA textural triangle is superimposed on the discriminant space.

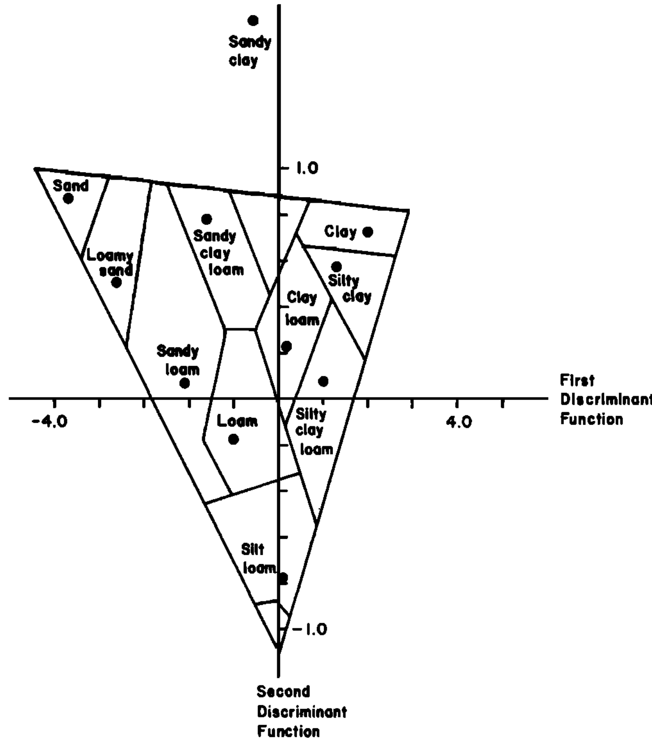

Fig. 8. Plot of the two most important discriminant function scores for the eleven group analysis. A distorted USDA textural triangle is superimposed on the discriminant space.

Discussion

As in the previous studies of this data set we found that of all the physical soil descriptors available, variability in texture was most closely related to variability in the soil moisture parameters. In previous work this result led to a simple tabulation of the statistical properties of the parameters in each textural class, a useful step in understanding the variability of the parameters. We have been able to extend these results in two ways. The discriminant analyses suggest an intuitive qualitative explanation for the observed relationship between parameter distribution and soil textural characteristics. The regression analyses provide a quantitative means of predicting the expected statistical properties of the parameters for a given soil texture.

Soil textural classes are determined uniquely by a combination of three variables, the percent sand, silt, and clay content of the soil. In this system, there are in reality only two independent variables, and these variables define a planar space such that each textural class occupies a unique region of the space. The discriminant analyses on the hydraulic parameters resulted in two important functions, each of which produces a single variable that is a linear combination of the hydraulic parameters. These two functions are orthogonal and can also be taken to define a planar space which may be divided into unique regions. The striking result of this analysis was that the two spaces showed a definite one-to-one mapping. That is, for a "typical" soil of a given textural class, the sand-silt-clay space is isomorphic with the hydraulic parameter space. It is intuitively reasonable that the hydraulic characteristics of a soil are determined by the particle size distribution of the soil. It would also seem reasonable that any set of hydraulic parameters that can define a planar space which provides the same discrimination between soil samples as a planar space based on the particle size distribution would be the minimum set of hydraulic parameters necessary to characterize the hydraulic behavior of the soil (at least to the same degree of resolution as that provided by the particle sizes). Thus we can infer that the parameters studied in this paper provide a nearly complete description of the hydraulic characteristics of soils given the information available.

Of more practical importance are the results of the regression analyses. The fact that the variances as well as the means of the hydraulic parameters are functions of soil textural class has not been reported before. That there is more inherent variability in the parameters in certain classes is perhaps not surprising. That the variability can be explained so simply as a univariate function of the sand, silt, or clay content is surprising. The large reductions in F ratios from the analysis of variance (see Figure 6) suggest that the regression equations are very robust since they can remove so much of the pattern in the parameter distributions. It must be emphasized that the patterns extracted in this analysis, while significant, are still embedded in a large amount of noise. The parameter variances for each textural class are not small relative to the means (see Table 3), and the patterns we observed may have been detectable only because of the large data set available for analysis. For any particular soil sample or small group of samples, the relationships described above may be obscured.

Attempts to model the observed spatial variability of soil moisture are commonly based on an assumed variance in the moisture parameters for a given soil type. Reliable estimates of the size of the variance to be used (or for that matter of the parameter means) have been lacking. Furthermore, the manner in which these means and variances might change in heterogeneous systems of mixed soil types has not been investigated either. The results presented here, having been derived from a large, diverse set of soil samples, should be indicative of the true pattern of variability in the hydraulic parameters. The use of these parameter class means and standard deviations for a known soil textural type may improve the predictions from stochastic models utilizing a homogeneous soil. The use of the regression equations for the parameter means and standard deviations should add increased sophistication to models which incorporate distinct layers of different soil textures. Because the regressions are continuous in the variables, it may be possible to construct models that are based on continuous spatial variation in physical soil properties (such as sand or clay content) which provide even better simulations of soil moisture. For all cases, knowing the patterns of parameter variability will greatly reduce the dimensionality of the modeling problem and increase the realism of the results.

Acknowledgment. This research was supported by the U.S. Army Research Office.

References

Arya, L. M., and J. F. Paris, A physical model to predict the soil moisture characteristic from particle size distribution and bulk density, *Soil Sci Soc. Am. J.*, *45*, 1023–1030, 1981.

Brakensiek, D. L., Comments on "Empirical equations for some soil hydraulic properties" by Roger B. Clapp and George M. Hornberger, *Water Resour. Res.*, *15*, 989–990, 1979.

Brakensiek, D. L., R. L. Engleman, and W. J. Rawls, Variation within texture classes of soil water parameters, *Trans. ASAE*, *24*: 335–339, 1981.

Brooks, R. H., and A T. Corey, Hydraulic properties of porous media, *Hydrol. Pap. 3*, 27 pp., Colo. State Univ., Fort Collins, 1964.

Campbell, G. S., A simple method for determining unsaturated conductivity from moisture retention data, *Soil Sci.*, 117, 311–314, 1974.

Childs, E. C., *The Physical Basis of Soil Water Phenomena*, 493 pp., Wiley-Interscience, New York, 1969.

Clapp, R. B., and G. M. Hornberger, Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, *14*, 601–604, 1978.

Cooley, W. W., and P. R. Lohnes, *Multivariate Data Analysis*, John Wiley, New York, 1971.

Ghosh, R K., Estimation of soil moisture characteristics from mechanical properties of soils, *Soil Sci.*, *130*, 60–83, 1980.

Gupta, S. C., and W. E. Larson, Estimating soil water retention characteristics from particle size distribution, organic matter percent and bulk density, *Water Resour. Res.*, *15*, 1633–1635, 1979.

Holtan, H. N., C. B. England, G. P. Lawless, and G. A. Schumaker, Moisture-tension data for selected soils on experimental watersheds, *Rep. ARS 41-144*, 609 pp., Agric. Res. Serv., Beltsville, Md., 1968.

Kendall, M. G., and A. Stuart, *The Advanced Theory of Statistics*, vol. III, 557 pp., Halfner, New York, 1968.

McCuen, R. H., W J. Rawls, and D. L. Brakensiek, Statistical analysis of the Brooks-Corey and the Green-Ampt parameters across soil textures, *Water Resour. Res.*, *17*, 1005–1013, 1981.

Nakano, M., Pore volume distribution and curve of water content versus suction of porous body, 1, Two boundary drying curves, *Soil Sci.*, *122*, 5–13, 1976.

Nie, N. H., C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, *Statistical Package for the Social Sciences*, McGraw-Hill, New York, 1975.

Rawls, W. J., and D. L. Brakensiek, Estimating soil water retention from soil properties, *J. Irrigat. Drain. Div. Am. Soc. Civ. Eng. 108*, 166–171, 1982.

Rawls, W., P. Yates, and L. Asmussen, Calibration of selected infiltration equations for the Georgia Coastal Plain, *Rep. USDA-ARS-S-113*, 110 pp., Agric. Res. Serv., Beltsville, Md., 1976.

Selvin, H. C., and A. Stuart, Data-dredging procedures in survey analysis, *Am. Statist.*, *20*, 20–23, 1966.

Tatsuoka, M. M., *Multivariate Analysis*, John Wiley, New York, 1971.

U.S. Department of Agriculture, Soil survey manual, *U.S. Dep. Agric. Agric. Handbk.*, *18*, 503 pp., 1951.

R. B. Clapp, B. J. Cosby, T. R. Ginn, and G. M. Hornberger, Department of Environment Sciences, Clark Hall, University of Virginia, Charlottesville, VA 22903.

(Received September 16, 1983;
revised January 28, 1984;
accepted February 6, 1984.)