SPATIAL VARIABILITY OF HYDRAULIC PROPERTIES IN A MULTI-LAYERED SOIL PROFILE

Dirk Mallants¹, Binayak P. Mohanty², Diederik Jacques¹, and Jan Feyen¹

Unsaturated hydraulic properties of field soils are needed for water flow and solute transport calculations in variably saturated soils. The purpose of this study was to quantify the spatial variability and spatial cross-correlation of estimated parameter values of a flexible retention model that was fitted to measured retention data. Moisture retention characteristic (MRC) curves were measured on 100-cm² undisturbed soil cores collected at 180 locations along a 31-m-long transect in a three-layered soil profile at depths of 0.1, 0.5, and 0.9 m. Sampling intervals in the horizontal direction were alternately 0.1 m and 0.9 m. Saturated hydraulic conductivity \( (K) \) was determined on the same soil cores using a constant head permeameter. The drying part of the MRC curves was described by the four-parameter retention model of van Genuchten with fitting parameters, namely \( \theta_0, \theta_r, \alpha, \) and \( n \). Spatial variability of the five parameters, \( \theta_0, \theta_r, \alpha, \) and \( n \), and \( K \), was investigated for the three horizons using conventional statistics and geostatistical techniques. Maximum coefficient of variation \( (CV) \) was found for \( K \), i.e., 599%, 322%, and 897% for the 0.1-, 0.5-, and 0.9-m soil depths, respectively. Macropores and small sampling volume contributed to this large variability of \( K \). When all three soil depths are considered, residual water content \( (\theta_r) \) and shape factor \( \alpha \) showed moderate heterogeneity with a maximum \( CV \) of 156 and 53%, respectively. Small spatial heterogeneity was observed for shape factor \( n \) and saturated water content \( \theta_0 \), with a maximum \( CV \) of 22 (for 0.1-m depth) and 8% (for 0.9-m depth), respectively. Most hydraulic parameters at different layers exhibited convex experimental semivariograms that could be described by means of spherical models, with a spatial range between 4 and 7 m. The correlation scales for cross-semivariograms for pairs of cross-correlated parameters were found to be of similar magnitude to those pertaining to the direct semivariograms of correlated variables.

HYDROLOGICAL and geological processes are known to vary in space (Nielsen et al. 1973; Delhomme 1979). The spatial heterogeneity of the processes can be characterized statistically in terms of the probability density function (pdf) (Warrick et al. 1977). Pure statistical treatment of the heterogeneity problem, however, ignores the existence of spatial correlation. A geostatistical analysis of the spatial heterogeneity, on the other hand, calculates correlations between the observations made at different neighboring locations. The observed correlation structures can then be used in practical applications such as estimating values of the selected properties at unsampled locations by means of kriging (Webster and Burgess 1980) or co-kriging techniques (Vauclin et al. 1983) or can be used in designing sampling networks. They can also be used in the generation of synthetic data or parameter fields, which are required for stochastic flow and transport modeling (Ababou 1988; Hopmans et al. 1988).

Different studies (e.g., Gajem et al. 1981; Russo and Bresler 1981a; Jury 1985; Ünlü et al.
1990; Mohanty et al. 1994a) have exhibited different spatial correlation structures for soil hydraulic properties such as saturated/unsaturated hydraulic conductivity, saturated and residual soil water content, sorptivity, and pore-size distribution parameter. The effects of the spatial variability of soil physical properties on water and solute distributions in field soils were investigated by a number of soil scientists. Based on a numerical analysis with spatially correlated hydraulic properties, Russo and Bresler (1981b) demonstrated that the state variables—such as pressure head, wetting front position, and salt concentrations—can also be characterized by a certain correlation distance. In a recent numerical study, Tseng and Jury (1993) demonstrated the effect of different correlation lengths and parameter variabilities on field-scale hydrologic behavior and the implications for in situ measuring of the hydraulic conductivity. The effect of spatial variability in \( K_s \) on observed solute breakthrough in undisturbed soil columns was reported by Mallants et al. (1995a).

All of these earlier studies indicated the importance and the need for proper characterization of the spatial variability of soil hydraulic parameters, especially for the field-scale applications of flow and transport models.

This study investigates the spatial variability of the soil–water retention parameters and the saturated hydraulic conductivity measured for three pedogenetic soil layers of a 31-m-long transect. First, the probability density function for each soil hydraulic parameter for each soil horizon will be estimated. Second, spatial dependence will be examined for the soil hydraulic parameters by constructing both direct semivariograms and cross-semivariograms for correlated random variables for the three pedogenetic layers (horizons) of the soil profile. The observed spatial correlation structures will be used in a future numerical study on field-scale flow and transport in the multi-layered soil.

MATERIALS AND METHODS

Soil Sampling

In a field experiment, three overlying soil horizons were sampled for a distance of 31 m. The transect was located in Bokkevoort, near Leuven, Belgium. The soil was classified as a well-drained sandy loam (Udalfluent or Eutric Regosol). Three horizons were identified in the first 100 cm; Ap (0–25 cm) and C1 (25–55 cm), which both developed on colluvial material, and C2 (55–100 cm), which was a textural B horizon.

In each horizon, 60 undisturbed 5.1-cm-long \( \times \) 5-cm-diameter soil cores were collected using a Uhland3 core sampler with an alternating sampling distance of 0.1 and 0.9 m. This spacing allowed the analysis of short-range variation in addition to large-range variation. For each of the 180 soil cores, we determined saturated hydraulic conductivity, \( K_s \), using a constant head permeameter (Klute 1965), bulk density, \( \rho_b \), and the drying part of the moisture retention curve (MRC) using a standard desorption technique (Hillel 1980). Water contents, \( \theta \) (cm3 cm-3), were measured in the 0–4.2 pF (pF = \( \log_{10} (|\psi|) \)), \( \psi \) is soil water pressure) range by means of a sand-box apparatus (for soil water pressures of 0.0, 0.5, 1.0, 1.5, and 2.0 pF) and pressure cells (for soil water pressures of 2.3, 2.8, 3.4, and 4.2 pF). Volumetric water content measurements were made using the undisturbed soil cores up to a pressure of 2.8 pF. At soil water pressure of 3.4 and 4.2 pF, the soil was removed from the core, wetted, and thoroughly mixed. The wet soil material thus obtained was put in 1-cm-high and 3.5-cm-diameter polyvinyl chloride (PVC) rings on pressure plates and allowed to equilibrate.

Data Analysis

Water retention data, \( \theta – \psi \), and bulk density were analyzed in terms of statistical moments of the empirical distribution functions. The van Genuchten model (1980) was fitted to each of the 180 observed moisture retention characteristic curves using the parameter optimization code RETC (van Genuchten et al. 1991). The parameter \( \theta – \psi \) model described by van Genuchten (1980) can be written as:

\[
\theta(\psi) = \theta_r + \frac{\theta_s - \theta_r}{(1 + (\alpha|\psi|)^n)^m}
\]

where the subscripts \( r \) and \( s \) denote residual and saturated water content, respectively, \( \alpha \) is the inverse of the air entry value (cm-1), \( n \) represents the slope of the MRC curve, and \( m = 1 - 1/n \).

General statistical parameters were calculated for the three sets (for three soil layers) of hydraulic properties \( \theta_i, \theta_s, \alpha_i, n_i, K_s_i \) for \( i = 1, 2, \ldots, 60 \). In a previous study, Mallants et al. (1995b) applied a series of seven transformations to the raw \( K_s \) and van Genuchten parameters, followed by the Shapiro-Wilk test-statistic (Shapiro and Wilk 1965) for normality. P-values for transformed parameters that best corresponded to a normal distribution were subse-

\footnote{Mention of trade names does not constitute endorsement of the manufacturing company.}
sequently used to generate a multi-variate normal pdf. This study, however, only considers a logarithmic transformation if the raw data were not Gaussian distributed. In this way, the comparison of the parameters' spatial structure between different horizons is not obscured by the fact that different transformations occur for the same parameter. Also, interpretation of the semivariograms and cross-semivariograms is easier if simple transformations are used to normalize the data.

The semivariogram is one of the major tools used for analysis and modeling of the spatial variability of regionalized variables. A basic assumption that is used in semivariogram analysis is that a regionalized variable \( Z(x) \) is said to be stationary of order 2 (second-order stationarity). This requires the expected value of \( Z(x) \) to be independent of location \( x \), and the covariance for each pair of variables \( [Z(x_1), Z(x_2)] \) exists, and depends only on \( h \), the separation vector. The second criterion also implies that the variance of \( Z(x) \), \( \sigma^2 \), is stationary, and is thus independent of \( x \). A slightly more general set of conditions is obtained for the intrinsic hypothesis (Matheron 1963), where the stationarity assumptions are expressed in terms of the difference \( [Z(x_1) - Z(x_2)] \) of the regionalized variable. The assumption of variance stationarity was checked by split-window technique as proposed by Mohanty et al. (1991). The semivariogram estimator, \( \gamma(h) \), was used as proposed by Matheron (1963):

\[
\gamma(h) = \frac{1}{2N(h)} \left\{ \sum_{i=1}^{N(h)} [z(x_i + h) - z(x_i)]^2 \right\}
\]

where \( N(h) \) is the number of pairs separated by the lag distance \( h \), and \( z(x_1), z(x_2), \ldots, z(x_n) \) are data values measured at spatial locations \( x_1, x_2, \ldots, x_n \). If two or more regionalized variables are spatially intercorrelated, the cross-semivariogram is given by Vauclin et al. (1983):

\[
\gamma_{12} = \gamma_{21} = (1/2) \quad E[z_1(x) - z_1(x+h)\|z_2(x) - z_2(x+h)]
\]

which can be estimated by

\[
\hat{\gamma}_{12} = \frac{1}{2N(h)} \left\{ \sum_{i=1}^{N(h)} [z_1(x_i) - z_1(x_i + h)\|z_2(x_i) - z_2(x_i + h)] \right\}
\]

Cross-semivariogram becomes useful when two or more variables that are cross-correlated have to be estimated using co-kriging at unsampled locations, where the primary variable is undersampled (because of higher level of difficulty and/or the cost involved) whereas the secondary variable (covariate) is far easier to determine or has been assembled routinely during surveys.

The spherical semivariogram model adopted in this paper to describe the spatial structures is defined as:

\[
\gamma(h) = \begin{cases} C_s & 0 < h \leq A \\ C_s + C_c & h > A \end{cases}
\]

where \( C_s \) is the nugget effect (represents small-scale structure that occurs at distances smaller than the sampling interval, microheterogeneity, and experimental error), \( C_c \) is the structural component, \( A \) is the range of spatial dependence, and \( C_s + C_c \) is the sill (total variance). Procedures for estimation of the parameters of the parametric semivariogram include, among others, least squares analysis, eye-balling, and use of kriging by the jack-knifing procedure (Vauclin et al. 1983). The jack-knifing method uses kriging to estimate the value for an observation at measurement location \( i \) based on the \( N-1 \) remaining observations. Validation criteria for this technique include the kriged average error (KAE), the kriged reduced mean square error (KRMSE), and the kriged mean square error (KMSE) (Springer and Cundy 1987). The KAE parameter is defined as:

\[
KAE = \frac{1}{N} \sum_{i=1}^{N} (z_i - \hat{z}_i)
\]

where \( z_i \) and \( \hat{z}_i \) are, respectively, the observed and estimated value at location \( i \) based on the \( N-1 \) remaining observations. The optimum value of KAE is zero or, in other words, there is no systematic error. The second criterion (KRMSE) expresses the consistency between kriged errors \( (z_i - \hat{z}_i) \) and the standard deviation of the observations, \( s \) as:

\[
KRMSE = \left( \frac{1}{N} \sum_{i=1}^{N} \frac{(z_i - \hat{z}_i)}{s} \right)^{1/2}
\]

The optimum value for KRMSE is one. The accuracy of the estimation is given by the kriged mean square error (KMSE), which has an optimal value of zero. It is defined as:

\[
KMSE = \left( \frac{1}{N} \sum_{i=1}^{N} (z_i - \hat{z}_i)^2 \right)^{1/2}
\]
These criteria were optimized by iteratively adjusting the variogram parameters in Eqs. (5) and (6). Calculation of the experimental semivariogram, the subsequent fitting of the theoretical model, and the validation of the model parameters by solving the kriging equations were done using the computer code of Englund and Sparks (1988).

RESULTS AND DISCUSSION

Statistical Analysis

Table 1 lists the major statistical parameters for the soil water content (θ_{34} is volumetric water content at pressure pF 3.4, etc.) and the bulk density for each soil horizon, assuming all data come from the same population and are independent. In general, retention data showed a greater CV at pressure heads greater than pF 2.3 in comparison with retention data at smaller pF values. Similar observations were reported by Greminger et al. (1985) for a field drainage experiment and by Shouse et al. (1995) for laboratory data obtained from small soil cores. The increased variability of retention data at drier conditions has important implications for sampling density (to be shown later). For θ_{3,4} and θ_{4,2} in the Ap horizon, this behavior is only weakly present. At soil-water pressures of 3.4 and 4.2 pF, variances for water contents for all horizons are very small. This opposite trend of the variance in comparison with the CV is expected because at 3.4 and 4.2 pF, water is released by more uniform pore sizes. Note the rather small values for skewness and kurtosis coefficients, indicating that for most pressure heads the data can be approximated by a normal pdf. The small difference between mean and median values also suggests a normal distribution. An exception is the C1 horizon, which displays much higher values for the skewness and kurtosis, such that a lognormal pdf is more appropriate. This is confirmed further by the relatively small W values (W-statistic for normality, Shapiro and Wilk 1965) for the C1 in comparison with the Ap and C2 (the closer is the W value to 1 the higher the chance the data follow a normal pdf). Figure 1 shows the spatial heterogeneity along the transect for bulk density and water content at 2.3 pF in

| Table 1 |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Statistical analysis of retention data |
| Ap horizon |
| | ρ₀ (g/cm³) | θ₀ | θ₀,5 | θ₁ | θ₁,5 | θ₂ | θ₂,3 | θ₂,8 | θ₃,4 | θ₄,2 |
| Mean | 1.42 | 0.423 | 0.406 | 0.404 | 0.398 | 0.351 | 0.289 | 0.217 | 0.091 | 0.073 |
| Median | 1.41 | 0.422 | 0.404 | 0.404 | 0.395 | 0.350 | 0.284 | 0.215 | 0.090 | 0.072 |
| Variance | 0.006 | 0.0009 | 0.0008 | 0.0008 | 0.0008 | 0.0007 | 0.0014 | 0.0025 | 0.00005 | 0.00005 |
| CV (%) | 5.5 | 7.1 | 7.0 | 7.0 | 7.3 | 7.4 | 13.0 | 23.0 | 7.9 | 9.8 |
| Skewness | -0.775 | 0.175 | 0.244 | 0.323 | 0.170 | 0.398 | -0.228 | 0.164 | 2.24 | 0.585 |
| Kurtosis | 1.881 | -0.678 | -0.327 | -0.289 | -0.527 | 0.222 | 1.590 | 0.147 | 11.503 | -0.055 |
| W* | 0.95 | 0.97 | 0.98 | 0.98 | 0.98 | 0.97 | 0.97 | 0.98 | 0.85 | 0.96 |

C1 horizon

| Mean | 1.54 | 0.382 | 0.361 | 0.353 | 0.328 | 0.295 | 0.258 | 0.206 | 0.101 | 0.075 |
| Median | 1.54 | 0.380 | 0.360 | 0.350 | 0.325 | 0.293 | 0.254 | 0.202 | 0.101 | 0.074 |
| Variance | 0.002 | 0.0004 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0004 | 0.0007 | 0.00009 |
| CV (%) | 2.6 | 5.5 | 4.7 | 4.7 | 5.2 | 6.3 | 8.3 | 16.7 | 8.2 | 13.0 |
| Skewness | 0.163 | 0.446 | 1.454 | 2.159 | 4.196 | 4.861 | 2.230 | 1.027 | -0.111 | 0.159 |
| Kurtosis | -0.219 | -0.108 | 5.819 | 9.287 | 24.483 | 29.174 | 9.096 | 4.086 | 0.92 | -0.491 |
| W* | 0.98 | 0.97 | 0.91 | 0.86 | 0.66 | 0.55 | 0.85 | 0.94 | 0.99* | 0.97 |

C2 horizon

| Mean | 1.53 | 0.460 | 0.437 | 0.429 | 0.411 | 0.386 | 0.357 | 0.281 | 0.111 | 0.078 |
| Median | 1.53 | 0.457 | 0.436 | 0.430 | 0.411 | 0.385 | 0.354 | 0.286 | 0.108 | 0.077 |
| Variance | 0.002 | 0.0012 | 0.001 | 0.001 | 0.0008 | 0.0006 | 0.0005 | 0.0009 | 0.0001 | 0.0001 |
| CV (%) | 2.9 | 7.4 | 7.3 | 7.2 | 7.1 | 6.5 | 6.4 | 10.7 | 11.0 | 14.6 |
| Skewness | -0.15 | 0.254 | 0.335 | 0.384 | 0.424 | 0.387 | 0.191 | -0.014 | 1.106 | 0.257 |
| Kurtosis | -0.677 | -0.662 | -0.771 | -0.712 | -0.610 | -0.598 | -0.314 | -0.735 | 1.331 | 2.196 |
| W* | 0.96 | 0.96 | 0.95 | 0.94 | 0.95 | 0.98 | 0.98 | 0.96 | 0.92 | 0.93 |

*a* Normality test using the Shapiro-Wilk test statistic.

*b* Normally distributed at a confidence level >0.90.
the three horizons (Ap, C1, C2). From Table 1 we note the surface horizon clearly has a lower bulk density in comparison with the deeper layers, most probably a result of large amount of macro pores, higher root density, and higher organic carbon content in Ap versus C1 and C2. These factors also explain the higher variability of bulk density (1.155 to 1.558) and water content at 2.3 pF (0.174 to 0.39) in the Ap horizon.

The number of samples, N, necessary to estimate the mean with a given accuracy, d, for a given probability level (Steel and Torrie 1980) is given as:

\[ N = \frac{u^2 \sigma^2}{d^2} \]

(10)

where \( \sigma \) is the standard deviation, and \( u \) is the degree of confidence (%). The underlying assumptions for Eq. (10) are (i) the samples are independent and (ii) the samples are drawn from a normal distribution. The first assumption is met for those hydraulic parameters that are not spatially correlated. For the parameters exhibiting spatial dependency (to be discussed later), the overall sample variance \( \sigma^2 \) may be underestimated, and the number of samples estimated from Eq. (10) should be considered as a first approximation. Based on the statistical parameters from Table 1, we further assume that our retention data are normally distributed. Table 2 lists the sample numbers required to estimate the mean value of a given parameter within \( \pm 10\% \) of the true value at a 0.05 significance level (95% of the time). As could be expected from the low CV at and near saturation, very few samples \( N = 2 \) (Ap), 1 (C1), 2 (C2) are required for an accurate estimation of the mean water content in the pressure range 0–2 pF. However, when the soil becomes drier, more samples are needed to accurately estimate the mean water content. For instance, at a pressure of 2.8 pF, 10 times more samples are needed for the Ap \( (N = 20) \) and C1 \( (N = 10) \) horizons, and two times more are needed for the C2 horizon \( (N = 4) \). When the pressure head increases further to 4.2 pF, once again, lesser samples are needed in Ap and C1 horizons. The C2 horizon, however, showed opposite behavior, nevertheless, considerably more samples are required compared with the values in the wet range. Notice that three different apparatuses were used to determine the water content at different soil water pressure heads, which may or may not have influenced the heterogeneity in the data. The tendency of the variability of soil water content to increase as the soil be-
comes drier has also been observed by Warrick and Nielsen (1980) and Shouse et al. (1995). From these values, we concluded that the number of samples collected in this study was large enough to characterize the mean value of the WRC parameters within ±10% of the true mean (95% confidence level). On the other hand, this is clearly not the case for the $K_s$ parameter. As expected, the $K_s$ parameter requires more observations than the MRC parameters. For example, Fig. 2 shows that estimating the mean value of $K_s$ (for the Ap) within ±10% of the true mean for 95% of the time requires a sample size of approximately 13,000 observations. The required sample numbers are comparable to those reported by Anderson and Cassel (1986). In Table 2 we note further that to estimate the mean $K_s$, requires up to three times fewer observations in Ap than in C1 and C2 horizons.

Statistical moments for the estimated moisture retention model parameters (Eq. 1) and the measured $K_s$ parameter are presented in Table 3. For the lognormally distributed parameters, mean ($\mu$), variance ($\sigma^2$), and coefficient of variation (CV) for the back-transformed parameters were calculated from:

$$\mu = \exp(\mu_{ln} + 0.5\sigma^2_{ln})$$  \hspace{1cm} (11)

$$\sigma^2 = \exp(\sigma^2_{ln} + 2\mu_{ln})(\exp(\sigma^2_{ln}) - 1)$$  \hspace{1cm} (12)

$$CV = \frac{\exp(\sigma^2_{ln}) - 1}{\exp(\sigma^2_{ln})}^{1/2}$$  \hspace{1cm} (13)

where $\mu_{ln}$ and $\sigma^2_{ln}$ are the mean and variance for the log$_e$-transformed variable.

Of note in Table 3 is the extremely large heterogeneity for $K_s$. The CV for $K_s$ ranges from 322% for the C1 horizon to 897% for the C2 horizon. This variability is much larger than the values reported by Jury (1985) based on a compilation of field data. However, Anderson and Cassel (1986) observed a CV of 3300% for their $K_s$ measurements in the A horizon of a Portsmouth sandy loam. This large variability was attributed to the remains of channels of tree roots. The large variability of $K_s$ in our sandy loam soil is presumably caused by the presence of macropores in combination with a small sample volume (100 cm$^3$). Bypass flow through large pores was demonstrated using methylene blue dye on 20 × 20-cm saturated soil columns (Mallants et al. 1994). Moreover, the variability of $K_s$ increases as sample volume decreases (Baker 1982; Lauren et al. 1988; Mohanty et al. 1994b). In the study of Mohanty et al. (1994b), $K_s$ variability was highest when measured with a constant head laboratory permeameter in a comparative study with four in situ methods, presumably because of the presence of large pores.
or absence of open-ended macropores. When macropores are present in the soil, the isolation of short soil cores tends to induce artificial boundary conditions. As a result, water will flow through macropores that are continuous over the whole length of the core, and hence the large values for the mean and the variance. Similar observations were reported by Anderson and Bouma (1973) for $K$, measurements on soil cores of different lengths and by Lauren et al. (1988) for $K$, measurements on attached (in situ) and detached (isolated) soil columns.

The effect of sample volume on variability of $K$, was addressed by Mallants et al. (1995b) who measured $K$, on three different soil volumes with increasing diameter (i.e., 0.05, 0.20, and 0.30 m) and length (i.e., 0.051, 0.20, and 1.0 m). The variance of log$_e$-transformed $K$, decreased with increasing sample size, from 3.7 to 1.74 and 0.74, respectively. Thus, the sample size used represents the largest variability, which, in turn, needs a greater number of samples. Although a higher sample volume will be more representative of field-scale flow and transport processes, collection of large columns may be impractical for routine investigations because of the increased sampling efforts involved.

The variability of the parameters $\theta_s$ and $\theta_n$ is moderate to low, with values of CV ranging from 54.9 to 156.4% for $\theta_s$ and 5.1 to 7.6% for $\theta_n$, respectively. The substantially higher CV for $\theta_s$ in the C1 was attributable to 34 $\theta_s$ values that were estimated to be zero, whereas the next lowest value was 0.0016. In general, these CV values are very similar to the results obtained by Russo and Bresler (1981a) and Shouse et al. (1995). Note the lower saturated water content for C1 in comparison with C2, although the bulk densities are similar. This is presumably due to the higher clay content in the C2, i.e., 21.8% vs. 16.6% in the C1.

The shape parameters $\alpha$ and $n$ are moderately heterogeneous, with CVs ranging from 45 to 53% for $\alpha$ and from 11 to 22% for $n$, respectively. Moreover, differences in variability between the horizons are small. These results compare fairly well to those reported by Russo and Bresler (1981a) and Anderson and Cassel (1986).

The degree of spatial variability of each of the four moisture retention parameters together with log$_e$(K) along the transect is depicted in Fig. 3. The estimated residual water content $\theta_s$ for the C1 horizon shows many zero values. Values for the saturated water content ($\theta_n$) are lowest for C1 and are of equal magnitude in Ap and C2 horizons. This can be explained by the higher bulk density (see Table 1) as a result of the reduced root activity in the C1 horizon. Higher $\theta_s$ values in C2 compared with those in C1 are the result of a higher clay content in the C2 horizon, i.e., 21.8% vs. 16.6% in the C1. The $\alpha$ parameter (inverse of air entry value) shows highest variability in the C2 and lowest in the Ap horizon. Variability in $\alpha$ can thus be interpreted as heterogeneity of pore sizes. The shape parameter $n$ clearly displays less heterogeneity in the C1 horizon in comparison with the Ap and C2. The smaller mean $n$ parameter for C1 (1.386) in comparison with the mean values for Ap and C2 (1.754 and 1.788) corresponds to a
Fig. 3. Spatial variability of estimated ($\theta_r$, $\theta_s$, $\alpha$, $n$) and measured ($K_r$) hydraulic parameters for Ap (left), C1 (middle), and C2 (right) horizon.

less steep retention curve. In other words, on average, the rate of decrease of water content with increasing soil-water pressure is smallest for C1. The $K_r$ parameter shows a very distinct variability in all horizons (Fig. 3), even at the smallest sampling interval of 10 cm. Table 3 shows the CV for $K_r$ in the C2 is more than twice as large as $K_r$ in the C1. The difference in variability in $K_r$ may be explained by a nonuniform redistribution of migrating clay particles in combination with the presence of macropores. From field observations and methylene blue experiments, we found that macropores are present in depths of up to 1 m or more.

Stationarity of the Variance

Stationarity of the variance was inspected by plotting of the interquartile-range-squared, IQ^2
vs. the median for local neighborhoods. These neighborhoods were defined by dividing the transect into 12 equal-sized windows, each having five observations. For each window, we computed the IQ$^2$ and the median for the raw data. In a second step we calculated the same statistical parameters using the Gaussian transformed data. The variance stabilizing effect of the transformations is demonstrated in Fig. 4 for the hydraulic parameters $\theta$, and $n$ pertaining to the C1. Also shown are the pdfs for the raw and log$_e$-transformed data. The results in Fig. 4 are for the C1 horizon, but similar observations were made in the Ap and C2 horizons (results not shown). It is evident from Fig. 4a and e, for raw data, as the median increases, the measure of variation (IQ$^2$) also increases, indicating nonstationarity of the variance. The correlation coefficient for the IQ$^2$–median pairs is 0.29 for the raw $\theta$, data (Fig. 4a), whereas for the log$_e$ $\theta$, (Fig. 4b) it is 0.11. For the $n$ parameter, the scatterplots (Fig. 4e) show that for the raw data, IQ$^2$ and median pairs are roughly proportional to one another, with a correlation coefficient of 0.63. However, in the log$_e$ scatterplot (Fig. 4f), the IQ$^2$ and median pairs are nearly uncorrelated ($r = 0.18$). If the data are transformed to a nearly normal distribution (see Fig. 4c and d for $\theta$, and 4g and h for $n$), the variance has become fairly homogeneous. Similar variance stabilizing effects were obtained by Mohanty et al. (1991) for their saturated hydraulic conductivity data using a log$_e$ transformation.

Semivariogram Analysis

For each of the four estimated retention parameters as well as for the measured $K_s$, parameter experimental semivariograms were computed for the three horizons. The semivariograms for the parameters $\theta$, $n$, and $K_s$ for all three horizons were based on log$_e$-transformed data. Because unreliable values for $\gamma$ (h) are obtained when $N(h) < 50$ (Journel and Huijbregts 1978), lag distances were chosen such that $N(h)$ is always larger than 50. Figure 5 shows experimental semivariograms for all five parameters pertaining to the three horizons. For the Ap, spatial structure was found for log$_e$ $\theta$, $\alpha$, and log$_e$ $K_s$, whereas $\theta$, and log$_e$ $n$ displayed random variation (pure nugget). Results for the C1 horizon indicated spatial structure for log$_e$ $\theta$, $\alpha$, and log$_e$ $n$ and random variation for $\theta$, and log$_e$ $K_s$. The C2 horizon displayed spatial structure for $\theta$, log$_e$ $\theta$, $\alpha$, and log$_e$ $n$ and random variation for log$_e$ $K_s$. Because the same transformations were used across different soil horizons, comparison of the parameter spatial dependency across horizons can be readily made. In general, the range of spatial correlation does not differ much between horizons and is almost always less than 5 m (except for log$_e$ $n$ in the C1). Most parameters in the C1 horizon exhibit the lowest overall variance (sill value) in comparison with the other two layers, except for the $\alpha$ parameter which shows the highest overall variability.

The spherical model (Eqs. (5) and (6)) was fitted to $\gamma(h)$ for parameters that showed a clear spatial dependence. Optimal model parameters were obtained by using the jack-knifing cross-validation procedure previously defined, resulting in values for KAE, KRMSE, and KMSE sufficiently close to their optimum. An exception is the log$_e$ $\theta$, parameter with values for KRMSE considerably larger than the optimum value of one, indicating that the kriged errors are larger than the standard.
deviation of the observations. The best-fit models are shown in Fig. 5. “Nugget” models (= overall sample variance) are also shown for those parameters exhibiting only random variation. Optimized semivariogram model parameters (Table 4) revealed that the nugget component \( C_n \) ranged from 26% (C2 horizon) to 71% (C1 horizon) of the sill value \( (C_s + C_n) \). In general, in the Ap and C1 the nugget effect \( C_n \) dominates, whereas in the C2 the structural component \( C_s \) dominates the spatial variability of different hydraulic parameters. At least for the water retention parameters, the high nugget effect is probably not attributable to microheterogeneity at distances smaller than 0.1 m, which was our smallest sampling distance. Rather, it is more likely caused by experimental uncertainties and/or the presence of non-optimal parameter estimates.

The log\( \theta \), \( \alpha \), and log\( n \) parameters pertaining to the Ap horizon showed spatial dependency over a distance of approximately 4 m (Fig. 5). The \( \theta \), and log\( n \) parameters, however, displayed random variation. The spherical model describes the experimental semivariograms fairly well, with the cross-validation criteria in general close to their recommended optimum. Spatial structure in the C1 horizon was observed for log\( \theta \), \( \alpha \), and log\( n \) (Fig. 5). All other parameters showed a pure nugget behavior. Better spatial dependence of log\( K \) in the Ap horizon in comparison with that of C1 is presumably attributable to higher biological activity in this horizon, which is most
likely to be concentrated around tree roots. Because trees were present at both sides of the entire transect length, at a distance of approximately 1 m, their influence may have been substantial. Spatial variation of root density along the transect was not investigated here.

The textural B layer (C2 horizon) showed spatially correlated parameters over a distance of 4 (log$_e$ $\alpha$, $\alpha$, and log$_e$$n$) to 4.5 m ($\theta_0$). The approximate linear behavior of the experimental semivariogram for $\theta_0$ (open circles in Fig. 5) indicated the presence of a trend. After removing the linear trend, the semivariogram was recalculated based on the residuals (solid circles). For $K_n$, no spatial dependency was observed. Again, validation criteria for the estimated spherical models are fairly close to their optimum (Table 4).

The log$_e$ $K_n$ parameter is spatially correlated only in the Ap. It has been shown by Lauren et al. (1988) that measurements of $K_n$ in soils with macropores can be highly affected by the sample volume, especially if the sample does not contain a representative amount of macropores. In other words, at least for the C1 and C2 horizons, the sample volume (100 cm$^3$) is probably smaller than a representative elementary volume (REV), and the spatial structure of $K_n$ if present, may go undetected.

The issue of whether or not our sampling volume encompasses a REV has been investigated based on soil structure units. According to Bouma (1985) a REV should contain approximately 20 elementary units of soil structure (ELUS). From soil survey information, we estimated the size of soil structure units to be 10–50 mm (medium to coarse subangular blocky (Soil Survey Staff 1975)) in the Ap horizon. An individual soil aggregate or ped has an approximate maximum volume of 125 cm$^3$, requiring a REV of approximately 2500 cm$^3$. This is much larger than our sampling volume and explains the large variability observed here.

Summarizing, for those parameters with a clear spatial structure, the range of spatial correlation is approximately the same in each horizon and equal to 4 m. This is an indication that several processes influencing spatial variability, work at the same scale, although the size of the transect itself may also influence the estimated range of spatial dependence.

**Cross-semivariogram Analysis**

Sample correlation coefficients were calculated for all five parameters for the three horizons (see Table 5). For all horizons, positive correlation was observed between $\alpha$ and log$_e$$\theta_0$, whereas log$_e$$n$
TABLE 5

Sample correlation coefficients between variables for each horizon

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \theta )</th>
<th>( \log(\theta) )</th>
<th>( \alpha )</th>
<th>( \log(n) )</th>
<th>( \log(K) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ap horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(\theta) )</td>
<td>-0.004</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>-0.140</td>
<td>0.389*</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(n) )</td>
<td>0.070</td>
<td>-0.193</td>
<td>-0.524*</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>( \log(K) )</td>
<td>-0.009</td>
<td>-0.008</td>
<td>-0.130</td>
<td>0.233</td>
<td>1.000</td>
</tr>
<tr>
<td>C1 horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(\theta) )</td>
<td>0.098</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>-0.12</td>
<td>0.52*</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(n) )</td>
<td>0.32*</td>
<td>0.08</td>
<td>-0.54*</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>( \log(K) )</td>
<td>-0.14</td>
<td>-0.11</td>
<td>-0.04</td>
<td>0.030</td>
<td>1.000</td>
</tr>
<tr>
<td>C2 horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta )</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(\theta) )</td>
<td>-0.140</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.123</td>
<td>0.510*</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(n) )</td>
<td>-0.120</td>
<td>-0.411*</td>
<td>-0.870*</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>( \log(K) )</td>
<td>0.096</td>
<td>0.060</td>
<td>0.144</td>
<td>-0.003</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*Significant at the 0.05 probability level.

and \( \alpha \) showed a negative correlation. Correlation was also found between \( \log(n) \) and \( \log(K) \), for \( \theta \) (for \( \text{Ap} \)), \( \log(n) \) and \( \theta \) (for \( \text{C1} \)), and \( \log(n) \) and \( \log(\theta) \) (for \( \text{C2} \)). Only correlation coefficients significantly different from 0 at the 0.05 significance level were retained in the subsequent estimation of the cross-semivariograms using Eq. (4). Yates and Warrick (1987) suggested that co-kriging is superior to ordinary kriging for predictions for 0.5 or larger correlation coefficients between variables. Although several of our correlation coefficients were smaller than the recommended value, we pursued calculation of the cross-variograms because co-kriging was not our primary concern. Rather we wanted to illustrate the existence of spatially cross-correlated hydraulic parameters in view of future applications with two- or three-dimensional numerical flow models.

Typical examples of observed experimental cross-semivariograms along with the theoretical model are shown in Fig. 6 for the pairs \( \alpha-\log(\theta) \), \( \text{Ap,} \log(n)-\alpha \) (C1), and \( \log(n)-\alpha \) (C2). Figure 6 further reveals a range of spatial correlation for the parameter combination \( \log(n)-\alpha \) that is more than

TABLE 6

Cross-semivariogram parameters estimated for the spherical model:

\( C_0 \) nugget component, \( C \) spherical component, \( A \) range (m)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( C_0 )</th>
<th>( C )</th>
<th>( A ) (m)</th>
<th>( C_0/(C_0+C) ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ap horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha-\log(\theta) )</td>
<td>0.00005</td>
<td>0.0005</td>
<td>5</td>
<td>9.1</td>
</tr>
<tr>
<td>( \alpha-\log(n) )</td>
<td>-0.0023</td>
<td>-0.0002</td>
<td>4</td>
<td>53.5</td>
</tr>
<tr>
<td>( \log(n))( -\log(K) )</td>
<td>0.0000</td>
<td>0.0900</td>
<td>4</td>
<td>0.0</td>
</tr>
<tr>
<td>C1 horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha-\log(\theta) )</td>
<td>0.000045</td>
<td>0.000113</td>
<td>4</td>
<td>28.5</td>
</tr>
<tr>
<td>( \alpha-\log(n) )</td>
<td>-0.0002</td>
<td>-0.00022</td>
<td>10</td>
<td>47.6</td>
</tr>
<tr>
<td>( \log(n))( -\theta )</td>
<td>0.0003</td>
<td>0.0005</td>
<td>4</td>
<td>37.5</td>
</tr>
<tr>
<td>C2 horizon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha-\log(\theta) )</td>
<td>0.000015</td>
<td>0.000065</td>
<td>4</td>
<td>18.8</td>
</tr>
<tr>
<td>( \alpha-\log(n) )</td>
<td>-0.0005</td>
<td>-0.0003</td>
<td>4</td>
<td>14.3</td>
</tr>
<tr>
<td>( \log(n))( -\log(\theta) )</td>
<td>-0.0000</td>
<td>-0.006</td>
<td>4</td>
<td>0.0</td>
</tr>
</tbody>
</table>
two times larger in the C1 compared with the C2. The estimated cross-semivariogram parameters were obtained by using the same jack-knifing procedure as for the direct semivariogram. Estimated parameters revealed that structural component (C_s) dominates nugget (C_0) in most cases (Table 6).

The best-fit values for A parameter range from 4 m to 10 m (Table 6). In most cases, the correlation scales that were estimated for the direct semivariograms for the log(θ), α, and log(K) parameters (see Table 4) are very close to the values obtained for the cross-semivariograms (Table 6). The existence of direct semivariograms and cross-semivariograms with identical correlation scales (for specific hydraulic properties) was assumed by Mantoglou and Gelhar (1987) in the derivation of their stochastic unsaturated flow models. This purely theoretical assumption seems to be valid for at least a subset of our data and for the parameter combinations considered here.

The existence of several spatially cross-correlated hydraulic parameters has important consequences for modeling flow and transport with two- and three-dimensional numerical models. Such models are used together with probabilistic data-generation procedures, such as the Turning Bands Method (Tompson et al. 1989), and supply the model with the necessary two- and three-dimensional (cross-correlated) parameter random fields. Although such multivariate random field generators have been reported in the literature (e.g., Mantoglou 1987), few examples exist (e.g., Ababou 1988) where they have been used to explore the impact of spatial correlation structures of the hydraulic parameters on simulated field-scale water flow and solute transport in heterogeneous soils. Whether such an approach results in better predictions of field-scale hydrologic processes in comparison with using a single scaling factor set (e.g., Tseng and Jury 1993) requires additional investigation.

**SUMMARY AND CONCLUSION**

Spatial heterogeneity was examined for five soil hydraulic parameters (θ, α, n, and K) for three different depths in a layered soil profile. When the variability of the parameter values was expressed in terms of the coefficient of variation, three groups could be distinguished: (i) parameters that are extremely variable (K) with a maximum CV of 897%, (ii) moderately variable parameters (θ and α) with a maximum CV of 156%, and (iii) parameters with a weak variability (θ and n) showing a maximum CV of 27%. Statistical analysis showed that the sample size used in our study to estimate the mean within ±10% confidence interval at a 95% probability level were large enough for all parameters except K.

Prior to the geostatistical analysis, parameters showing a skewed distribution were transformed to a near-normal pdf. The transformations also guaranteed stationarity in the variance, as could be observed from interquartile-range-squared versus median plots. Non-stationarity in the mean was observed for the θ parameter in C2. Removal of a linear trend resulted in residuals with stationary mean. Experimental semivariograms showed spatial dependency for most parameters up to about 4 m. In the upper two horizons, microheterogeneity and/or experimental error dominated spatial structure (more than 60% of the overall variability was contributed by
nugget effect), whereas in the deepest layer, the opposite behavior was observed. Small scale variations in \(\log K\) values assumed the presence or absence of open-ended macropores. For the water retention parameters, microheterogeneity at distances less than 0.1 m is more unlikely to be the reason of the high nugget value such that experimental uncertainties may have been more important.

Cross-semivariograms were calculated for spatially cross-correlated pairs of hydraulic parameters. The estimated range of spatial correlation (\(4\)) for the theoretical spherical model varies from 4 to 10 m, which is very close to the values obtained for the direct semivariograms. These results suggest that the equivalence of the correlation scale for the direct- and cross-semivariogram may be valid for certain parameter combinations and certain field conditions.

REFERENCES


