Scale Effects in Estimating the Variogram and Implications for Soil Hydrology

Jon Olav Skøien* and Günter Blöschl

ABSTRACT

In monitoring and modeling landscape soil processes, the sampling and modeling scales should, ideally, be commensurate with the scales of the soil characteristics of interest. Unfortunately, this is usually not possible, both because the true covariance structure of the variable of interest is unknown a priori and because of logistical constraints. We examine the biases and random errors in variogram parameters that result from the choice of a sample scale triplet (spacing, extent, and support) that is not commensurate with the scales of the underlying "true" soil variability. We generate numerous two-dimensional random fields, from which we sample data sets randomly or on a grid. We use these samples to estimate the variogram parameters with various methods. The results indicate that both biases and random errors may be large and depend on the sampling scale triplet relative to the scale of the underlying soil variability. The maximum likelihood (ML) method of parameter estimation gives the smallest biases for largely spaced random samples, while the weighted least squares (WLS) method gives the smallest biases for largely spaced gridded samples. Nonparametric estimates exhibit smaller random errors but larger biases than estimates from the two parametric methods (ML, WLS).

V ADOSE ZONE–RELATED variables often exhibit some sort of organization at the landscape scale. In some cases it may be possible to explain the laws underlying the organization, but in other cases this may be very difficult. It may then be prudent to represent the soil-related variables as random variables (Webster, 2000). This variable can be soil moisture; concentrations of nutrients, pollutants, or bacteria; or dynamic properties related to any of these, such as hydraulic conductivity or denitrification rates. Assuming that the variable of interest can be considered a realization of a random field, geostatistical methods can be used both for interpolation of the variable between the measurements and for characterizing the spatial properties of the variable of interest in a parsimonious way by making use of the variogram. The sill of the variogram is a measure of the magnitude of the spatial variability, the correlation length is a measure of the degree of correlation in space (i.e., the scale of the underlying process), and the nugget is a measure of small-scale variability.

In this study we focused on the use of geostatistics for characterizing spatial properties in the landscape. Such characterization has been used in numerous scientific disciplines. Soil and hydrologically related studies include radon concentration in soils (Oliver and Khayrat,

2001), hydraulic conductivity in aquifers (Gelhar, 1993, p. 292), infiltration rates in agricultural landscapes (Haws et al., 2004), and soil moisture characteristics of landscapes (Mohanty et al., 2000; Western et al., 2004), among many other examples.

Common to all these studies is that the sampling scales are not necessarily commensurate with the scales of the underlying processes. The three sampling scales— spacing, extent, and support—were termed the *scale triplet* by Blöschl and Sivapalan (1995). In the two-dimensional case, as is of importance for environmental variables in landscapes, Blöschl (1999) defined them as characteristic length scales: spacing (L_S), extent (L_E), and support (L_A) (Fig. 1). When measuring, for example, soil moisture in an area of size A_{dom}, the spacing (L_S) is the average distance between the centers of each measurement and is a function of the size of the domain, A_{dom}, and the number of samples N. The extent (L_E) reflects the physical boundaries of the domain of measurements, and the support (L_A) is related to the support area A of the measurement, e.g., the surface area of a soil sample or the footprint of a pixel from remote sensing. In the case of a square, these relations are:

$$L_S = \sqrt{\frac{A_{dom}}{N}} \qquad [1]$$

$$L_E = \sqrt{A_{dom}} \qquad [2]$$

$$L_A = \sqrt{A} \qquad [3]$$

When the sampling scales are not commensurate with the scale of the underlying process, they will likely affect the estimated characteristics of the data. The data may then exhibit both biases and random errors that are related to the choice of sampling scale.

Large supports will typically result in underestimation of the variance and overestimation of the correlation length (Journel and Huijbregts, 1978). Romshoo (2004) analyzed remotely sensed soil moisture at different spatial scales and noted that increasing the support increased the estimated correlation length and decreased the variance. Russo and Jury (1987) found that large spacings led to overestimation of the correlation length because the small-scale variability will be aliased with larger scales (Nyquist, 1924). Beckie (1996) analyzed sampling scales associated with the spacing and support of a single groundwater field test and used filter functions in the frequency domain to estimate the variability between measurements, given their support and spacing. He found that increasing spacing increased the unresolved small-

Institute for Hydraulic and Water Resources Engineering, Vienna Univ. of Technology, Karlsplatz 13, A-1040 Vienna, Austria. Received 23 May 2005. *Corresponding author (skoien@hydro.tuwien.ac.at).
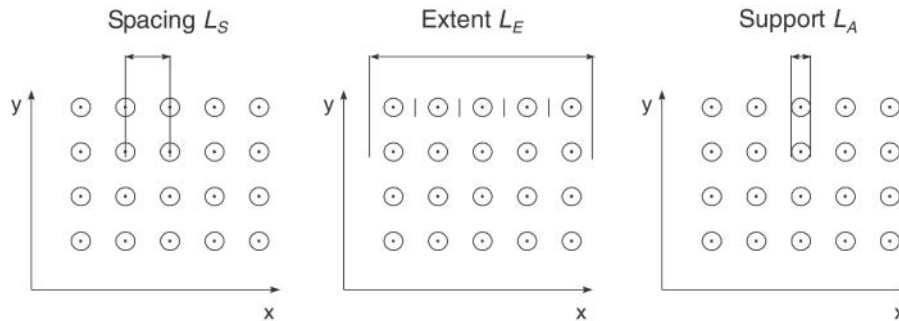
Fig. 1. The sampling scale triplet—spacing L_S, extent L_E, and support L_A—for the two-dimensional case.

scale variability, while increasing support decreased it. Gelhar (1993) compared estimates of the spatial correlation length of hydraulic conductivity with the extent of the analyzed domains, and suggested (Fig. 6.5 of Gelhar, 1993) that the correlation lengths tend to be on the order of 10% of the extent of the domain irrespective of the underlying hydrogeology of the aquifer. Similar results have been obtained by Blöschl (1999) for snow cover patterns. McBratney (1992) showed that the spatial variance of soil characteristics increases with extent. Dungan et al. (2002) discussed different scales used in spatial statistics with descriptions of their effect on different spatial characteristics. Western and Blöschl (1999) performed a more coherent analysis of spacing, extent, and support, using soil moisture data, and developed analytical expressions for the biases of variance and correlation length. They noted that the sampling scales affected the estimates of the variogram parameters depending on which scale of the scale triplet (spacing, extent, and support) is considered and provided a framework for reconciling the apparently conflicting results in the literature regarding scale effects on variogram parameters.

There have been several studies that examined uncertainties or random errors in the estimated variogram parameters. Cressie (1985) suggested that the uncertainty would depend on the variogram value and the number of samples in a bin. Shafer and Varljen (1990) proposed a jack-knifing method to estimate the confidence limits. Ortiz and Deutsch (2002) and Pardo-Igúzquiza and Dowd (2001) developed analytical expressions for describing the uncertainty of the estimates, taking the correlation between sample pairs into account, based on work by Menke (1989) and Woodbury and Sudicky (1991). Webster and Oliver (1992) examined the influence of the number of samples on the estimates of the local variogram for one particular experimental setup. They concluded that a variogram computed from 150 samples will often be satisfactory, while one derived from 225 samples will usually be reliable. Their findings were for one experimental setup, and estimation of the local variogram. In different applications the number of samples needed for reliably estimating the variogram may be different.

Sample variograms can be estimated in different ways (Matheron, 1965; Cressie and Hawkins, 1980; Omre, 1984), most of which are a variant of plotting the squared differences of the data values of pairs of points against lag and then averaging them within each lag bin. Fitting

a variogram model to the sample variogram can be done by, for example, ordinary least squares (OLS), generalized least squares, weighted least squares with approximations (WLS) (all from Cressie, 1991, p. 94–97), maximum likelihood (ML), and restricted maximum likelihood (REML) (Kitanidis, 1983; Kitanidis and Lane, 1985; Pardo-Igúzquiza, 1998). These methods assume that the type of variogram model is known. Gorsich and Genton (2000) proposed a method of finding the most likely type of variogram model by estimating the nonparametric derivative of the sample variogram values.

Zimmerman and Zimmerman (1991) compared seven estimators that can be used for estimating the parameters of a variogram. They also examined the estimates as a function of the parameters of the underlying variogram to get an indication of sampling scale effects. Their results showed that the performance of all methods was similar. They suggested that the loss of accuracy by using simple methods such as OLS and WLS is small compared with the extra computational burden of using more complex methods such as ML and REML. Cressie (1991, p. 100) discussed these results and questioned the robustness of the ML and REML methods, and related them to departures from the Gaussian assumption to which they are closely tied. Pardo-Igúzquiza (1998), however, suggested that the ML method is applicable to data that do not conform to a multivariate Gaussian distribution, although this may be difficult to verify in practice.

There is a distinction between the theoretical variogram, and the local variogram (Journel and Huijbregts, 1978, p. 192). The theoretical variogram represents all possible realizations of a process, while the local variogram is obtained by sampling exhaustively only within the area of interest. For a description of the general correlation structure of a variable the theoretical variogram is required, while for interpolation between observations in a study area, the local variogram is required. This paper focuses on how the sampling scale affects the estimated variogram parameters of the theoretical variogram. Specifically, the aim of this study was to identify the biases and random errors in variogram parameters that result from the choice of a sample scale triplet (spacing, extent, and support) that is not commensurate with the scales of the underlying "true" soil variability. We have chosen a simple correlation structure, assuming normally distributed values, stationarity, and absence of noise to focus on the base case. More complicated distributions are be-

yond the scope of this paper. The variogram parameters are estimated by the WLS and ML methods and are compared with the variogram parameters used for generating the random fields, along with the corresponding nonparametric estimates.

METHODS
Generating Random Fields and Sampling

We analyzed sampling scale effects by generating and analyzing a large number of random fields with a prescribed correlation structure. This provides an opportunity to analyze repeated realizations from a homogenously distributed variable with a known correlation structure.

The random field represents the spatial distribution of the variable of interest, and each realization represents one possible spatial distribution of such a variable. The random fields are assumed to be stationary; that is, the statistical characteristics of the population of the random field (such as the mean and the variance) were assumed not to change with spatial location. We also assumed that the univariate distribution of the random field conforms to a normal distribution with population mean μ and population variance σ^2. Finally, we assumed that the spatial correlation structure of the population of the random field could be represented by an exponential variogram:

$$\gamma(h) = c_0 + c_1\{1 - \exp[-(h/\lambda)]\} \quad [4]$$

where λ is the correlation length, h is the spatial distance between two points in the random field, and c_0 is the nugget effect, caused by microscale variability and measurement errors, while $c_s = c_0 + c_1$ is the sill of the variogram, equal to the variance σ^2 of the variable.

The variogram describes the variance between two points separated by the distance h. If no microscale variation or measurement errors are present, the nugget is zero. The correlation length is a measure of the average distance for which a variable is correlated in space. Small correlation lengths indicate that the variable varies erratically over short distances, while large correlation lengths indicate that the variable varies smoothly over short distances and variability starts to get significant at larger distances. The exponential variogram is consistent with a first-order autoregressive or Markov process (Webster and Oliver, 2001, p. 116) and hence is the simplest assumption one can make about the spatial variability of random fields.

The mean, variance, and correlation length of the underlying variogram are the true population characteristics. We assumed no nugget effect, and generated two-dimensional random fields with the prescribed statistical characteristics using the turning band method (TBM) (Mantoglou and Wilson, 1981, 1982) on a square grid of 1024 by 1024 points with grid size of Δx. We used 16 bands for generating the random fields. For each combination of the sampling scale triplet we generated 1000 realizations. All results in this paper are normalized with respect to the underlying random field. This corresponds to a random field with normal distribution $N(0,1)$ and a correlation length $\lambda = 1$. The realizations closely resembled the imposed correlation structure on average, although at small distances, the fields were a little too smooth. This will have a minor effect on some of the estimates, as discussed below.

Sampling

Throughout this paper we use dimensionless sampling scales L_S^*, L_E^*, and L_A^*, that is, the spacing, extent, and support scaled by the true correlation length λ of the population of the random field:

$$L_S^* = L_S/\lambda \quad [5]$$

$$L_E^* = L_E/\lambda \quad [6]$$

$$L_A^* = L_A/\lambda \quad [7]$$

In a first analysis we varied spacing and extent jointly. From each of the 1000 realizations for each extent, we took 16, 100, and 1024 samples. For a fixed number of samples, the spacing increases directly with the extent, similar to Eq. [1]:

$$L_S^* = L_E^*/\sqrt{N} \quad [8]$$

where N is the number of samples. For these analyzes, we took point samples from the grid ($L_A^* = 0$). In a second analysis we varied the support, fixing the extent as $L_E^* = 10$, with spacing being a function of the number of samples according to Eq. [8]. Again, we took 16, 100, and 1024 samples and aggregated an increasing number of points from the generated fields to represent the effect of increasing support. We assumed square support areas. In both analyzes we examined two different monitoring schemes, sampling on a regular square grid and random sampling.

In all instances with $L_A > L_S$, measurements will be partly overlapping, and the measurements closest to the border of the domain will also include values from outside what is defined as the extent of the domain. This can be the case in remote sensing, where the support of a pixel can be larger than the spacing between the pixels. We assumed that all sample values were error free. In a real world study, instrument errors will introduce additional uncertainty.

Estimation of Spatial Characteristics
Nonparametric Estimates

From the N samples $z(\mathbf{x}_i)$ at location $\mathbf{x}_i$ we estimated the sample mean ($\bar{z}$) and sample variance (s^2) as

$$\bar{z} = \frac{1}{N}\sum_{i=1}^{N} z(\mathbf{x}_i) \quad [9]$$

and

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}[z(\mathbf{x}_i) - \bar{z}]^2 \quad [10]$$

We estimated the sample variogram by the traditional estimator of Matheron (1965):

$$\hat{\gamma}(h) = \frac{1}{2n(h)}\sum_{i=1}^{n(h)}[z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})]^2 \quad [11]$$

where $h = |\mathbf{h}|$ is the spatial lag between two points. The summation over the number of pairs $n(h)$ is within bins, which we chose at logarithmical intervals. The integral scale of the population is defined as (Taylor, 1921; Russo and Jury, 1987; Western and Blöschl, 1999):

$$J = \int_0^\infty \left[1 - \frac{\gamma(h)}{\sigma^2}\right]dh \quad [12]$$

The sample integral scale we estimated from the sample variogram as:

$$\hat{J} = \sum_{i=1}^{N_0}\left[1 - \frac{\hat{\gamma}(h_{i-1}) + \hat{\gamma}(h_i)}{2s^2}\right]\Delta h_i \quad [13]$$

where N_0 is the bin number for which $\hat{\gamma}(h_{N_0}) = s^2$, that is, the bin number where the sample variogram first intersects a horizontal line $\hat{\gamma}(h) = s^2$. If the variogram reaches the sample variance between two bins, the exact intersection point is found by linear interpolation. We assumed $\hat{\gamma}(h_0) = 0$ and $\Delta h_i = h_i - h_{i-1}$. For the spatial correlation structure in Eq. [4] with no nugget effect, the theoretical integral scale is identical to the correlation length while for different correlation structures the integral scale can be smaller or larger than the correlation length, depending on the shape of the variogram. All nonparametric estimates are shown for comparison purposes with their parametric counterparts. More details on the nonparametric estimates are given in Skøien and Blöschl (2006).

Parametric Estimates

We estimated the variogram parameters by two different methods, assuming an exponential variogram model (Eq. [4]). First we used the WLS method of Cressie. This method optimizes the objective function Φ dependent on the set of parameters θ by:

$$\Phi(\theta) = \sum_{i=1}^{N_b} |N(h_i)| \left[\frac{\hat{\gamma}(h_i)}{\gamma(h_i|\theta)} - 1 \right]^2 \qquad [14]$$

where $N(h_i)$ is the number of pairs with separation distance h_i, N_b is the number of bins, $\hat{\gamma}(h_i)$ is the estimated sample variogram value for bin i, according to Eq. [11], and $\gamma(h_i|\theta)$ is the modeled variogram, given the parameter set θ. This method is simple and has two main advantages over OLS. First, bins with a higher number of pairs will get a higher weight. Second, the residuals close to $h = 0$, where the variogram values are small, will get higher weights because the errors are measured relative to the modeled variogram value. We examined three cases. First, we fitted a one-parameter model with no nugget effect ($c_0 = 0$) and assumed that the sill was equal to the sample variance s^2. Only the correlation length was fitted. In the second case, we still assumed no nugget effect, but we fitted the correlation length and the sill jointly. In the third case, we fitted the complete three-parameter model. We used the Shuffle Complex Evolution Method (Duan et al., 1992) for solving the optimization problem. This method needs limits for the parameter space. As we wanted to minimize the influence of these limits on the estimates, the possible parameter space had to include many orders of magnitude. It can in some cases be difficult to estimate the parameters in this case, especially if they are small. Therefore we logarithmized the parameters in the search procedure. This method makes it easier to find small parameter values, but at the same time increases the likeliness of estimating small parameters in cases where the optimum is flat.

The second method was the ML method, which maximizes the likelihood of actually having observed the data set given the parameters of the variogram model. This method is based on the assumption that the data conform to a multivariate Gaussian distribution. The joint probability density function of n experimental data can then be expressed as

$$p(\mathbf{z}|\theta) = (2\pi)^{-n/2} |\mathbf{V}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{z}-\mu)'\mathbf{V}^{-1}(\mathbf{z}-\mu) \right] \quad [15]$$

where $\mathbf{z}$ is the $n\times1$ vector of data, μ is the $n\times1$ vector of means, and $\mathbf{V}$ is the $n\times n$ variance-covariance matrix (Pardo-Igúzquiza, 1998). Separating the covariance matrix into a

variance and a correlation matrix, $\mathbf{V} = \sigma^2\mathbf{Q}$, the negative log likelihood function (NLLF) can be defined as:

$$L'(\hat{\beta},\hat{\sigma}^2,\theta|\mathbf{z}) = \frac{n}{2}[\ln(2\pi) + 1 - \ln(n)] + \frac{1}{2}\ln|\mathbf{Q}|$$
$$+ \frac{n}{2}\ln[(\mathbf{z} - \mathbf{X}\hat{\beta})'\mathbf{Q}^{-1}(\mathbf{z} - \mathbf{X}\hat{\beta})] \quad [16]$$

where $\hat{\beta}$ is the ML estimator of the drift parameters. The NLLF is minimized through an iteration procedure. Cressie (1991, p. 91) and Pardo-Igúzquiza (1998) give a more detailed description of the method, especially how to estimate the mean vector and the variance. We used a method implemented under the statistical environment R (R Development Core Team, 2004) by Ribeiro and Diggle (2001).

We used the ML method for examining the same three cases as with the WLS method (one-, two-, and three-parameter models). The computational demand of ML estimation rapidly increases with the number of samples. We have, therefore, only used this method to fit variogram parameters to 50 of the sample sets for each sampling scale for 1024 samples, while we fitted variogram parameters to all 1000 sample sets for 16 and 100 samples. The increasing demand of computational power is related to a particular advantage of ML estimation (Lark, 2000); that is, the ML method estimates the parameters directly from the cloud of squared differences of the variable between the points. Hence, no predefined lag classes are needed.

To reduce the effect of outliers and confine the search to parameters within a reasonable parameter space, we allowed the variogram parameters to vary within certain bounds for the WLS and ML methods. In the WLS method the estimated nugget was allowed to vary between zero and the estimated sample variance; in the ML method there was no upper limit for the nugget. In both methods, the estimated sill was allowed to vary between 0 and 10 times the estimated sample variance, and the estimated correlation length was allowed to vary between 0 and 10 for $L_E^* < 1$. For larger extents, the upper limit was set to 10 times the extent of the data set, the lower limit to $L_S^*/50$. The lower limit of the correlation length will affect the estimates of the correlation length, but in practice this is irrelevant since a fitted correlation length smaller than 2% of the spacing of the samples would not be acceptable.

Presentation of Results

From each realization for a given combination of the scale triplet, we estimated the nonparametric characteristics (mean, variance, and integral scale) by Eq. [9], [10], [13], and the variogram parameters by the WLS and ML methods. We chose the median as the representative value of all estimates because the distributions of most parameters were strongly skewed, and plotted the 25 and 75% quantiles as error bars. The error bars then represent the uncertainty of the estimates one encounters if only one sample set (i.e., one realization) is available. With 1000 realizations, the estimation error of the median is small. The error bars are only shown for every third sampling scale combination, and the estimates from the ML and WLS methods are slightly shifted for clarity of presentation.

Figure 2 shows an example of the sample variograms estimated from 10 realizations and samples from a 10 by 10 grid. The extent was $L_E^* = 5.6$, corresponding to a spacing of $L_S^* = 0.56$, and the support was $L_A^* = 0$. Although the extent is considerably larger than the correlation length ($\lambda^* = 1$), some of the sample variograms still appear nonstationary. Figure 2 also indicates the parametric estimates of sill (vertical dashed line) and correlation length (horizontal dashed line) from a two-parameter model, and the nonparametric estimates of sample variance (vertical solid line) and integral scale (hori-
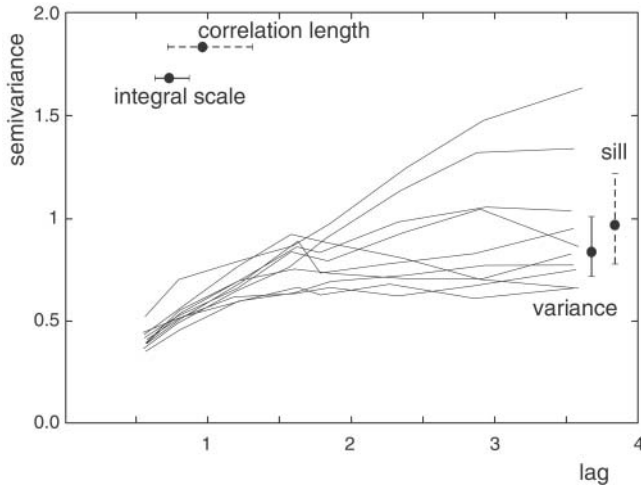
Fig. 2. Sample variograms obtained from 100 samples with extent $L_E^* = 5.6$ and support $L_A^* = 0$ from 10 realizations of the random field. Horizontal bars represent estimates of the integral scale (solid) and the correlation length (dashed) from a two-parameter model, and the vertical bars represent the variance (solid) and sill from a two-parameter model (median with 25 and 75% quantiles as error bars).

zontal solid line), together with their error bars, as calculated from the complete set of 1000 realizations.

We plotted the estimates in nondimensional form scaled by the (true) population characteristics. The nonparametric estimates of the spatial correlation structure (variance and integral scale) are

$$s^{2*} = \frac{s^2}{\sigma^2} \qquad [17]$$

$$\hat{J}^* = \frac{\hat{J}}{\lambda} \qquad [18]$$

and the parametric estimates (i.e., the parameters of the variogram model—nugget, sill, and correlation length) are

$$\hat{c}_0^* = \frac{\hat{c}_0}{\sigma^2} \qquad [19]$$

$$\hat{c}_s^* = \hat{c}_0^* + \hat{c}_1^* = \frac{\hat{c}_0 + \hat{c}_1}{\sigma^2} \qquad [20]$$

$$\hat{\lambda}^* = \frac{\hat{\lambda}}{\lambda} \qquad [21]$$

Since spacing and extent are related through the number of samples, we first present the results for spacing and extent together, followed by the results for the support.

RESULTS

Spacing and Extent

One-Parameter Model

Figure 3 shows the results when only correlation length has been fitted to the sample variograms from gridded samples. The green lines relate to the WLS method, the blue lines to the ML method. Unbiased estimates of the correlation length should produce medians of $\hat{\lambda}^* = 1$. For comparison, the red lines show the nonparametric estimates (i.e., the integral scale). For small extents, the WLS estimates are strongly negatively biased for all sample sizes. In these cases, the sample variograms increase monotonously with distance. The sill has been fixed equal to the sample variance, but the estimated correlation lengths increase with extent. The WLS estimates for small extents are almost equal to the estimates of the integral scale, but the uncertainty is slightly larger. The ML estimates for small extents are negatively biased for 16 samples, close to unbiased for 100 samples, and positively biased for 1024 samples. This is related to a tendency of the ML method of estimating the variogram considerably flatter than the WLS method for small extents, which is because the ML method appears to put more emphasis on the short distance part of the variogram and the TBM generated fields are slightly too smooth. Because the sill is fixed equal to the variance, the estimated correlation lengths for small extents are larger than those estimated by the two other methods. This effect increases with an increasing number of samples, and while the estimated correlation length is considerably underestimated for 16 samples, it is overestimated for 1000 samples. The small biases of the ML method for 100 samples are hence believed to be due to a compensation of two biases rather than an indication of a generally better performance than the other methods.

There is only a short range of sampling scales where both the WLS and the ML methods are able to produce
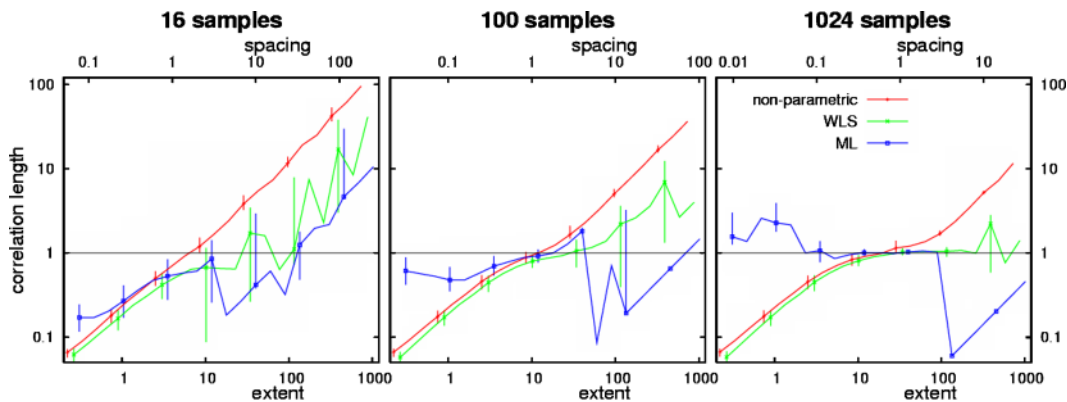

Fig. 3. Effect of spacing L_S^* and extent L_E^* on the estimated correlation length $\hat{\lambda}^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods, and integral scale $\hat{J}^*$ (red lines) for gridded sampling (median with 25 and 75% quantiles as error bars), for the one-parameter model.

close to unbiased estimates of the correlation length. Figure 3 indicates that the estimates are unbiased for $L_E^* > 8$ and $L_S^* < 3$. The range hence increases with an increasing number of samples. With 16 samples, the estimates are always biased. Unbiased estimates can only be achieved when the biases from small extent and large spacing cancel each other out. With 100 samples, the range of unbiased estimates is less than one order of magnitude for both the WLS and the ML methods. With 1024 samples, both methods can achieve unbiased estimates over more than one order of magnitude of the sampling scales relative to the scale of the underlying process.

For large spacings $L_S^* > 3$, the WLS estimates are positively biased. The estimates tend to fluctuate between the lower limit allowed and the spacing, and the uncertainty is large. In most cases the sample variograms do not exhibit any correlation (a steep increase to the sill at very short lags), so the WLS method will estimate the correlation length as a value between zero and the lag of the bin with the shortest lag, which has a lag close to the spacing. From the 1000 realizations, a large number of the estimates are equal to the lower limit. If this includes more than 500 of the estimates, the ensemble median is also equal to the lower limit. The ensemble medians from the ML method are close to the lower limit for large spacings. The parametric estimates of the correlation length for the WLS and ML methods are generally less biased than the nonparametric estimates (i.e., integral scale), but the uncertainty is much larger.

Figure 4 shows the results for random samples as opposed to the gridded samples in Fig. 3. When extent is small, the results are similar to the results for gridded sampling. This is because the biases and uncertainties are controlled by the extent in this case. For larger extents, and hence larger spacings, most estimates are less biased than the estimates from gridded samples. For 16 samples, the estimates from both the WLS and ML methods increase with increasing extent. With 100 samples, the scale range of unbiased estimates is significantly larger than for the gridded case (approximately between $L_E^* > 8$ and $L_S^* < 10$), but the ML method gives unbiased estimates for larger spacings than the WLS method. The

upper limit seems to increase slightly with an increasing number of samples. This is because there will be more sample pairs constraining the short part of the variogram, both increasing the resolution and the certainty of the estimates. The parametric estimates of the correlation length are again less biased than the nonparametric integral scale, and the uncertainties are larger. The ML method has the largest uncertainties. With 1024 samples the WLS and ML estimates of the correlation length are close to unbiased for a large range, and the uncertainty is small. This also indicates that the TBM method preserves the correlation structure of the underlying random field.

Two-Parameter Model

Figure 5 presents the results when sill and correlation length have been simultaneously fitted to the empirical variograms from gridded samples. For large extents the sill converges to unity, which indicates that the TBM method preserves the variance of the underlying random field. For small extents, the WLS method significantly overestimates both the sill and the correlation length. When extent is small, it is apparently not possible to separately estimate the sill and the correlation length. The derivative of the variogram is

$$\frac{d\gamma}{dh} = \frac{c_s}{\lambda}\exp(-h/\lambda)$$

When the extent is small, the exponential part will be close to 1, and the derivative becomes independent of the lag. The parameter estimation methods can then only identify the ratio between the sill and the correlation length, c_s/λ.

The ML estimates of the sill for small extent are negatively biased and close to the sample estimates of variance, but they are slightly less underestimated from 100 samples and 1024 samples. As the ML estimates of the sill are close to the estimated variance, the ML estimates of the correlation length are similar to the results from the one-parameter model (Fig. 3). For large extents, the sill is estimated close to the true sill, although for a small number of samples the uncertainty is larger
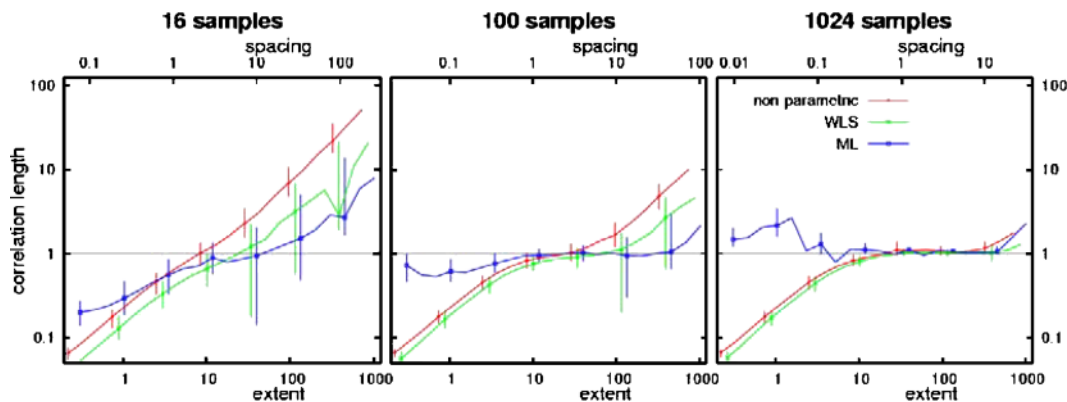

Fig. 4. Effect of spacing L_S^* and extent L_E^* on the estimated correlation length, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods, and integral scale $\hat{J}^*$ (red lines) for random sampling (median with 25 and 75% quantiles as error bars), for the one-parameter model.
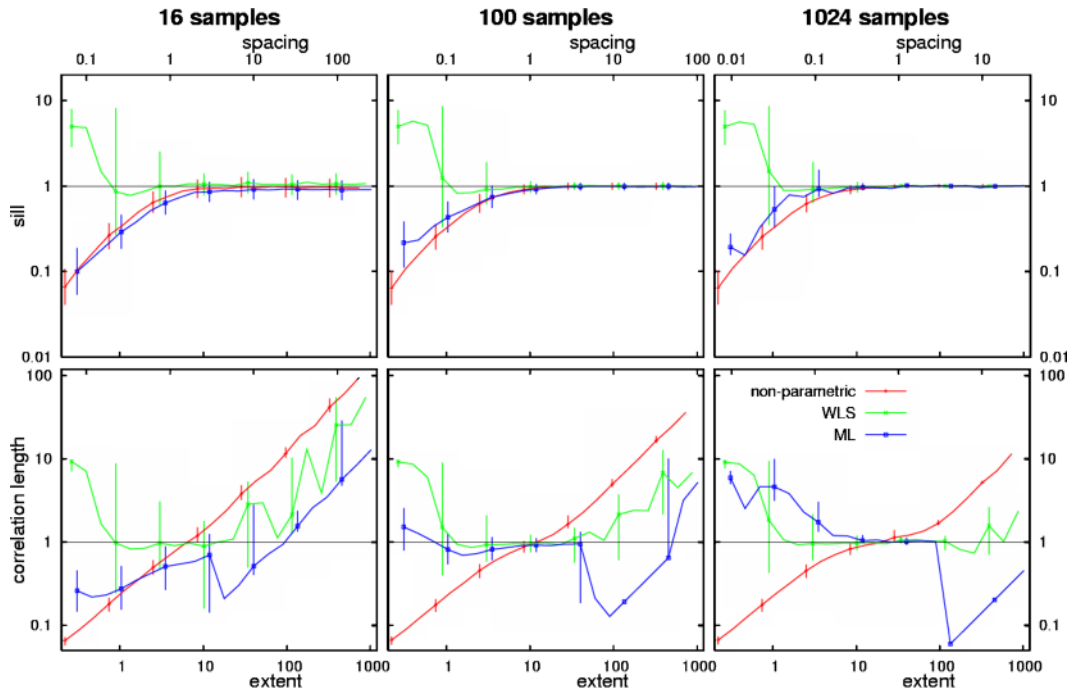
Fig. 5. (Top) Effect of spacing L_S^* and extent L_E^* on the estimated sill $\hat{c}_s^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods, and the sample variance s^{2*} (red). **(Bottom)** Effect of spacing L_S^* and extent L_E^* on the estimated correlation length $\hat{\lambda}^*$ (WLS green, ML blue) and the integral scale $\hat{J}^*$ (red). Gridded sampling (median with 25 and 75% quantiles as error bars). All are for the two-parameter model.

than that of the sample variance. For extents slightly smaller than 1, the estimated sill and correlation lengths from the WLS and ML methods are less biased than the estimated variance and integral scale, except the ML estimates from 16 samples. On the other hand, the uncertainty is much larger, even for the WLS estimates from 1024 samples. For large extents and spacings, the estimates of the sill are unbiased with small uncertainties, and the estimates of the correlation lengths are similar to the one-parameter model of Fig. 3. The estimates of correlation length are practically unbiased for $L_E^* > 2$ and $L_S^* < 3$. This means that the two-parameter model is able to provide unbiased estimates for smaller extents than the one-parameter model for which the limits were $L_E^* > 8$ and $L_S^* < 3$.

Figure 6 shows the same estimates as Fig. 5 but for a random sampling scheme. For extents $L_E^* < 5$, the sampling strategy does not appear to influence the estimates. For large extents and spacings, the figure shows the same effect as for a one-parameter model in Fig. 4 with a wider range of unbiased estimates than from gridded sampling.

Three-Parameter Model

Figure 7 shows the results of the three-parameter variogram model for gridded sampling. Unbiased estimates should produce median nuggets of $\hat{c}_0^* = 0$, median sills of $\hat{c}_s^* = 1$, and median correlation lengths of $\hat{\lambda}^* = 1$. Figure 7 (top) indicates that there is close to zero nugget effect for short spacings $L_S^* < 1$. For larger spacings, the estimated nugget increases for all sample sizes. The nugget estimated by the ML method is larger than that

of the WLS method and approaches the underlying variance for very large spacings. The estimates of the sill in the middle panels of Fig. 7 indicate that, for small spacings ($L_S^* < 1$), there are only small differences between this sill and the estimates from the two-parameter variogram model, as shown in Fig. 5. The estimates from 16 samples are always positively biased and are associated with larger uncertainties than the estimates from the two-parameter model. For large extents, the bias is small, but the uncertainties of the WLS and ML estimates are considerably higher than the uncertainties of the sample variance for 16 and 100 samples. In the case of 1024 samples, the uncertainties are only slightly larger than those of the sample variance. The ML method slightly underestimates the sill for 16 samples, while the WLS method slightly overestimates the sill. The ML estimates have significantly larger uncertainty than the ML estimates for 16 samples, and slightly larger uncertainties for estimates from 100 samples. Figure 7 (bottom) indicates that, for small extents, the estimates of the correlation lengths are similar to the two-parameter estimates shown in Fig. 5, except for a larger bias in the case of 16 samples. For large spacings $L_S^* > 3$, the estimates are positively biased, similar to the estimates from the one- and two-parameter model, but the uncertainties are larger. It appears to be impossible to fit a three-parameter variogram model to 16 samples with a reasonable level of accuracy. For 100 and 1024 samples, reasonable fits over a scale range of $L_E^* > 2$ and $L_S^* < 3$ are possible with the WLS and ML methods, but the uncertainty at the upper limit ($L_S^* = 3$) is much larger than for a two-parameter model.
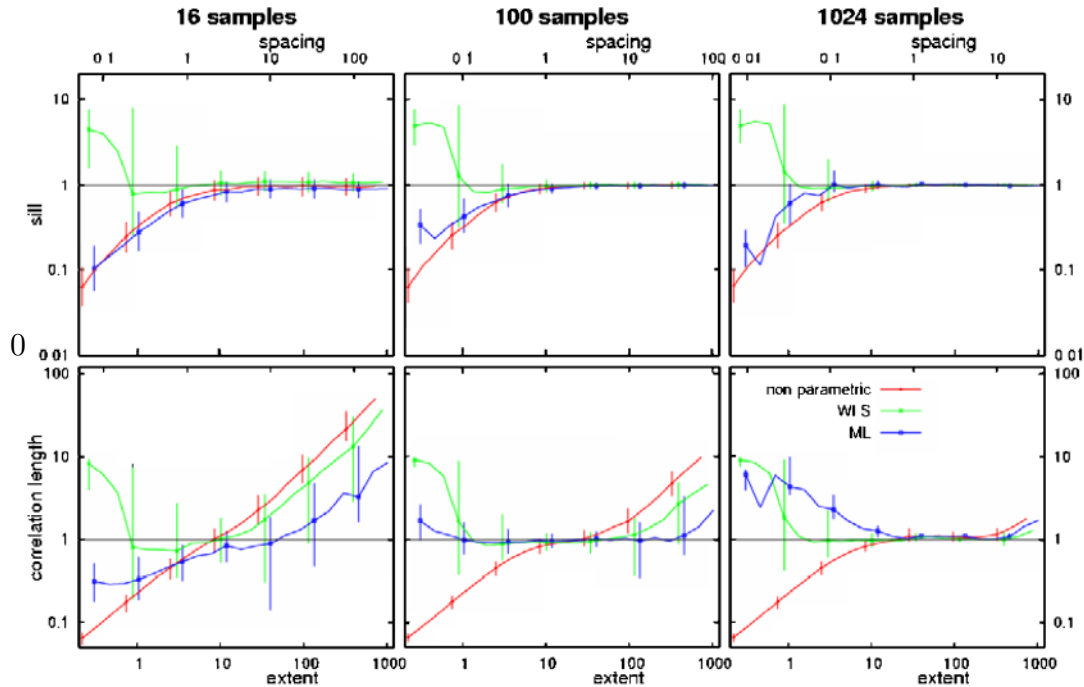
Fig. 6. (Top) Effect of spacing L_S^* and extent L_E^* on the estimated sill $\hat{c}_s^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods, and the sample variance s^{2*} (red). (Bottom) Effect of spacing L_S^* and extent L_E^* on the estimated correlation length $\hat{\lambda}^*$ (WLS green, ML blue) and the integral scale $\hat{J}^*$ (red). Random sampling (median with 25 and 75% quantiles as error bars). All are for the two-parameter model.

Figure 8 shows the same estimates as Fig. 7, but for a random sampling scheme. The results are generally similar to those in Fig. 7 but differ for large spacings. Figure 8 (top) indicates that the biases in the nugget first appear for spacings larger than in the gridded sampling case. Estimates of the sill are similar to the estimates from the gridded sampling scheme, but the correlation lengths can be estimated without bias over a wider range of sampling scales than for gridded sampling. For large spacings, the uncertainty appears to be slightly smaller than for gridded sampling. For large spacings, estimates of the correlation lengths from 100 and 1024 samples show similar biases as the integral scale, but the uncertainties are larger.

Support

One-Parameter Model

Figure 9 shows the results for the case where the support is large compared to the underlying correlation length, using a one-parameter model and gridded samples. The vertical line in each panel shows $L_S^* = L_A^*$. An extent of $L_E^* = 10$ has been used in all cases. The figure indicates that large supports generally increase the correlation length and the integral scale. This is related to the smoothing effects of aggregation, which reduces the variability over short distances. For 100 and 1024 samples, the estimates from the ML and WLS methods are quite similar as long as $L_A^* < L_S^*$ with little bias. For larger supports, the biases increase. The WLS estimates of the correlation length from 100 and 1024 samples are slightly less biased than the estimates of the integral scale. Both the WLS and the ML methods underestimate the correlation lengths for small supports from

16 samples, as the spacing is large relative to the underlying correlation length. The estimates in this case are similar to the results for the same spacing and extent in Fig. 3 ($L_E^* = 10$, $L_S^* = 2.5$).

The ML estimates exhibit significant biases for large supports. This is because the variograms tend to monotonously increase with lag if the support is large. This is similar to the situation on the left side of Fig. 3, where also the ratio between sill and correlation length is estimated, rather than the sill and correlation length separately. Similar to what was noted earlier, the ML procedure has a tendency of underestimating this ratio. For a fixed sill, the correlation length is overestimated, leading to significant biases for 100 samples, and even more so for 1024 samples as shown in Fig. 9.

Random sampling (not shown here) gave results similar to those shown in Fig. 9, with some differences for 16 samples, as spacings are also large in this case. For small supports, the biases and uncertainties are similar to the results for the same spacing and extent in Fig. 4.

Two-Parameter Model

Figure 10 shows the results for the case where the support is large compared with the underlying correlation length using a two-parameter variogram model and gridded sampling. The estimates of sill and variance (top panels in Fig. 10) decrease with increasing support, except for some estimates of the sill from the WLS method. The ML estimates of the sill are close to the estimates of the sample variance. For large supports, the variogram will appear nonstationary, and it is again only possible to estimate the ratio between sill and correlation length. The WLS estimates of the sill then increase. For yet larger
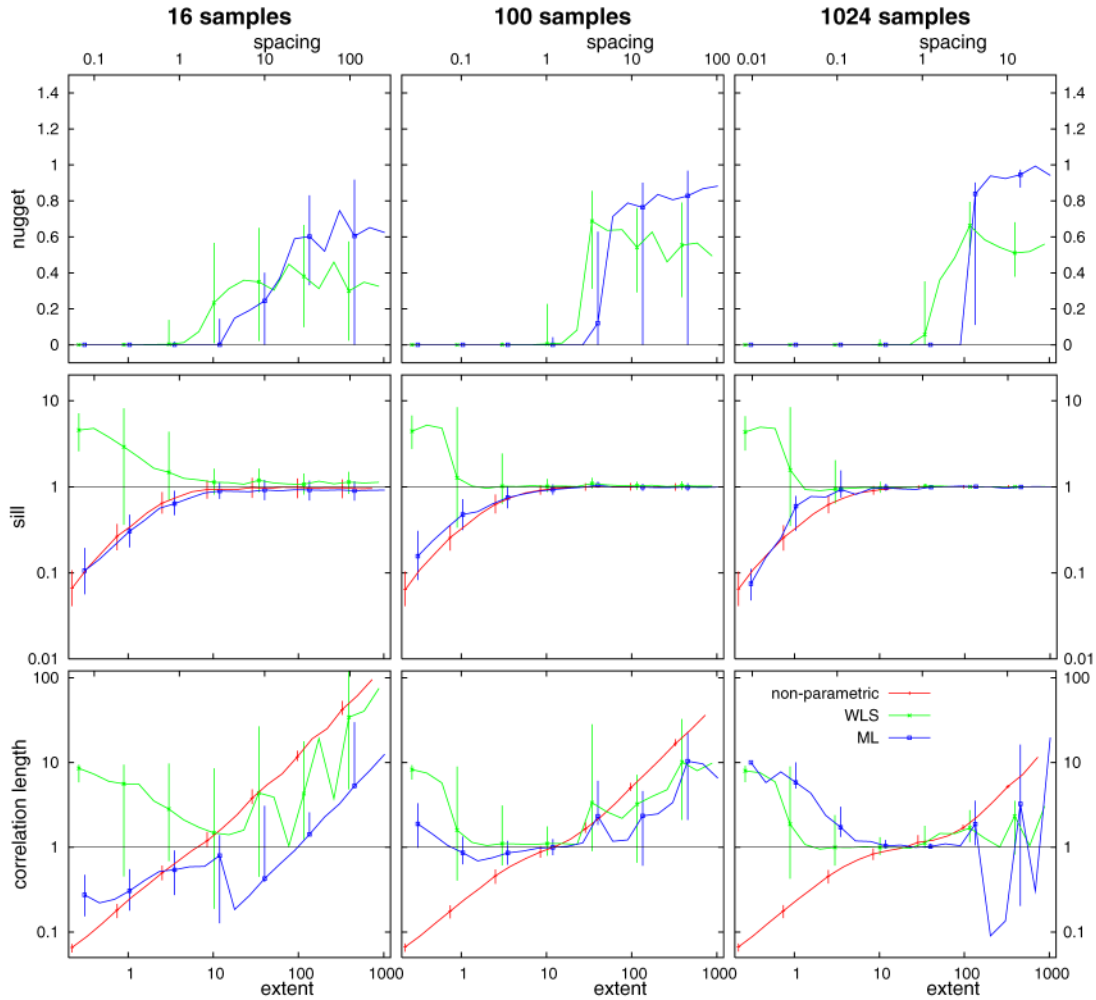
Fig. 7. (Top) Effect of spacing L_S^* and extent L_E^* on the estimated nugget $\hat{c}_0^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods. (Middle) Effect of spacing L_S^* and extent L_E^* on the estimated sill $\hat{c}_s^*$ (WLS green, ML blue) and the sample variance s^{2*} (red). (Bottom) Effect of spacing L_S^* and extent L_E^* on the estimated correlation length $\hat{\lambda}^*$ (WLS green, ML blue) and the integral scale $\hat{J}^*$ (red). Gridded sampling (median with 25 and 75% quantiles as error bars). All are for the three-parameter model.

supports, the effect of reduced sample variance due to smoothing reduces the estimated sill.

The second row of panels shows the estimates of correlation length and integral scale. For small supports, the results of all methods are similar to the results of the one-parameter model. For large supports, the ML estimates are still similar to the estimates from the one-parameter model. The estimates from the WLS method, however, increase much faster with increasing support than those from the one-parameter model, which is again related to the fact that only the ratios between sill and correlation length can be estimated in this case. The WLS method estimates larger sills than the ML method, similar to the case of small extents. For very large supports, the WLS method overestimates correlation lengths more significantly than the ML method.

Three-Parameter Model

Figure 11 shows similar results as Fig. 10 but for a three-parameter model. Biases in the nugget are only estimated in the case of 16 samples. This is because of the relatively large spacing of $L_S^* = 2.5$. The WLS estimates of the nugget have both higher median values and higher uncertainty than the ML estimates. For 100 samples and small support there is some uncertainty in the nugget, which again is related to the spacing ($L_S^* = 1$). The second row of panels of Fig. 11 shows the estimates of the sill. The results are similar to the results from the two-parameter model, except that the uncertainty is higher, especially for estimates from 16 samples. Figure 11 (bottom) shows the estimates of the correlation length together with the integral scale. For 100 and 1024 samples, the results are similar to the results of Fig. 10, except for small supports where the uncertainties are larger, especially for the WLS estimates. For 16 samples, the ML method estimates the correlation length close to the lower limit, while the WLS method strongly overestimates the correlation length.

Comparison with Expected Biases

Western and Blöschl (1999) presented analytical expressions for the expected variance and integral scale as
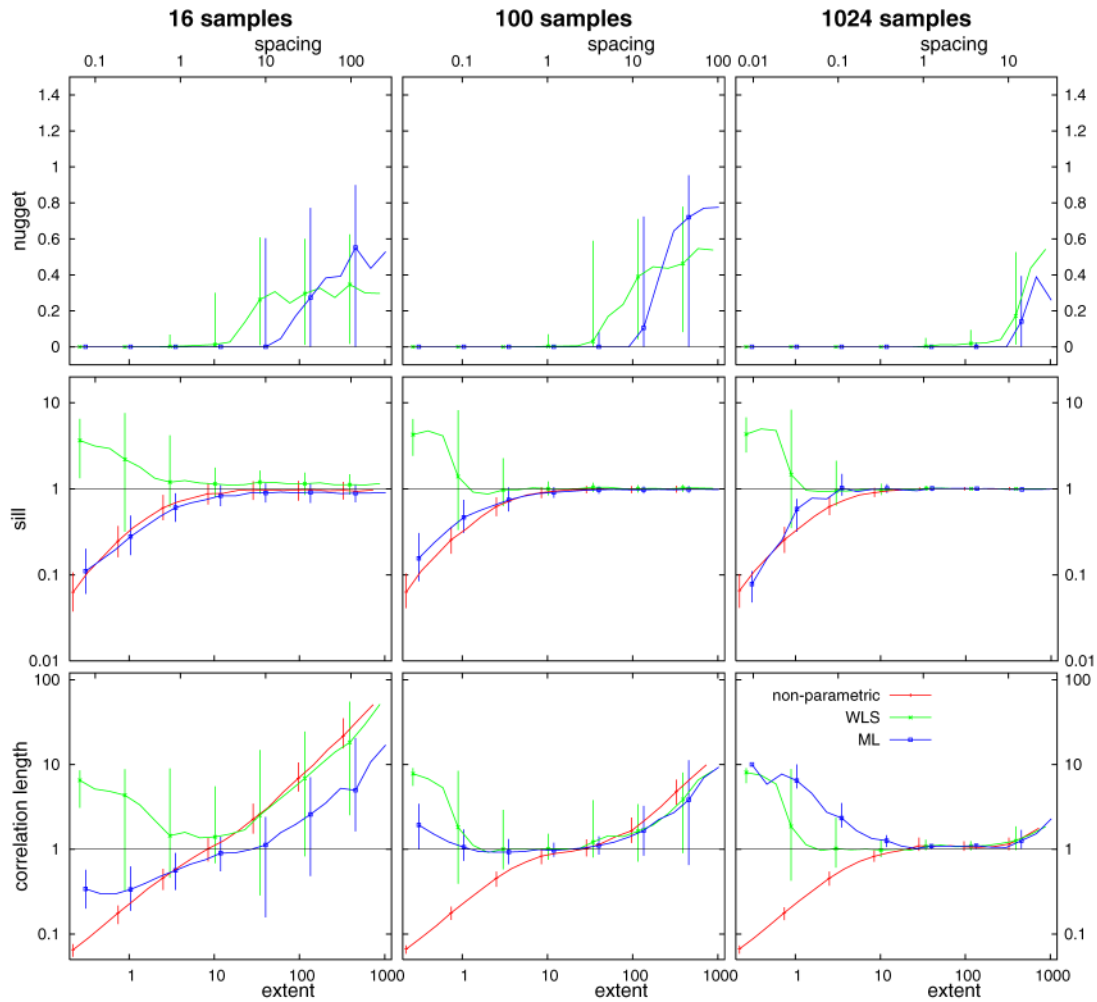
Fig. 8. (Top) Effect of spacing L_S^* and extent L_E^* on the estimated nugget $\hat{c}_0^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods. (Middle) Effect of spacing L_S^* and extent L_E^* on the estimated sill $\hat{c}_S^*$ (WLS green, ML blue) and the sample variance s^{2*} (red). (Bottom) Effect of spacing L_S^* and extent L_E^* on the estimated correlation length $\hat{\lambda}^*$ (WLS green, ML blue) and the integral scale $\hat{J}^*$ (red). Random sampling (median with 25 and 75% quantiles as error bars). All are for the three-parameter model.

a function of extent, spacing, and support. These are shown in Fig. 12a (thin black lines) together with the WLS estimates of the variogram parameters of a two-parameter model for 100 gridded samples. Figure 12b shows the corresponding values for the ML estimates. The estimates of integral scale in both figures are close

to the analytical expectations, while the parametric estimates of the correlation lengths differ from these expectations for both the WLS and ML methods. The summary also nicely demonstrates that depending on what component of the scale triplet is changed, the effect on the sill and the correlation length will differ.
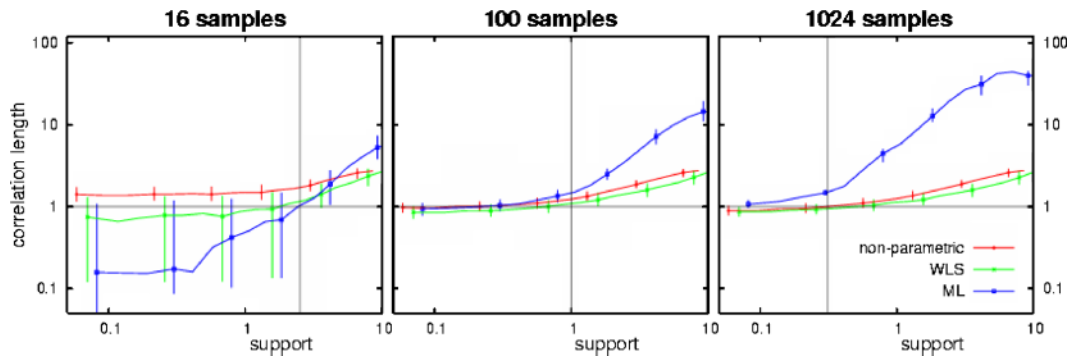

Fig. 9. Effect of support L_A^* on the estimated correlation length $\hat{\lambda}^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods, and integral scale $\hat{J}^*$ (red) for gridded sampling (median with 25 and 75% quantiles as error bars), for the one-parameter model. Vertical line in each panel shows $L_S^* = L_A^*$. $L_E^* = 10$, spacing L_S^* according to Eq. [8].
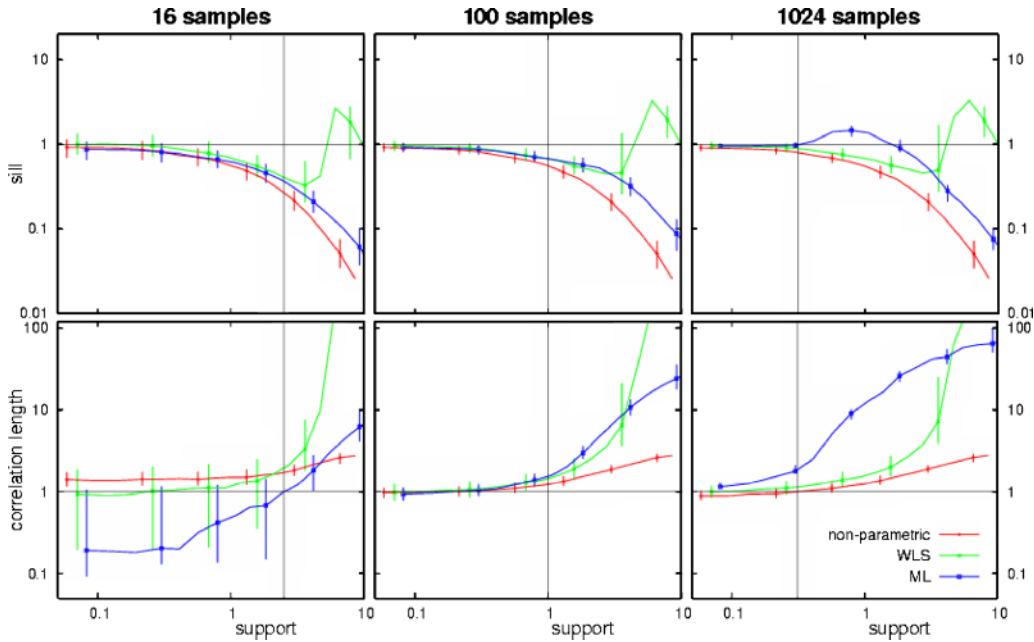
Fig. 10. (Top) Effect of support L_A^* on the estimated sill $\hat{c}_S^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods, and the sample variance s^{2*} (red). (Bottom) Effect of support L_A^* on the estimated correlation length $\hat{\lambda}^*$ (WLS green, ML blue) and the integral scale $\hat{J}^*$ (red). Gridded sampling (median with 25 and 75% quantiles as error bars). All are for the two-parameter model. $L_E^* = 10$, spacing L_S^* according to Eq. [8].

DISCUSSION AND CONCLUSIONS

There are two purposes of estimating a sample variogram and fitting a variogram model to the sample variogram. The first is to understand and characterize the spatial correlation of the variable of interest for which a theoretical variogram of the population of the random field is needed. The second purpose is interpolating the variable of interest within the study area to points where measurements are unavailable, for which a local variogram may be sufficient. This paper focuses on the first application.

The random variable examined in this paper can be any variable of interest in soil hydrology. The most typical variables are soil moisture, different kinds of concentrations, such as concentration of nutrients, or physical properties of the soil, such as its thickness, the thickness of pedological layers, soil density, and texture. In addition, it is also possible to use geostatistical methods to characterize spatially the dynamical properties of the soil, such as the hydraulic conductivities.

These simulations indicate that sampling scales significantly influence the results by biasing the estimates. This is not surprising because similar patterns were found in analyses of real data, such as those of Gelhar (1993) and Blöschl (1999), who showed that the estimated correlation lengths tended to be a function of the extent of the data. The parametric estimates are found to be less biased than their nonparametric counterparts, which is consistent with the findings of Russo and Jury (1987). While in many cases the biases of the parametric estimates are smaller than those of the nonparametric methods, their uncertainties are generally higher, especially for estimates of the correlation length. This uncertainty also increases with an increasing number of parameters. The number of samples is important in this context, as would be expected. An increasing number of samples reduce the uncertainty considerably. It is quite clear that correlation lengths cannot be estimated with any degree of confidence from 16 samples, since either the spacing is too large or the extent is too small. The window of unbiased estimates where the sampling scales are commensurate with the scale of the underlying variability is nonexistent in the case of 16 samples. For 100 samples, the variogram is based on almost 5000 sample pairs, but these might be highly correlated. This is described in methods of assessing the uncertainty of variograms (Ortiz and Deutsch, 2002; Pardo-Igúzquiza and Dowd, 2001). Because of the correlation, variogram parameters estimated from 100 points may still be highly uncertain. Webster and Oliver (1992) suggested that the number of samples should be at least 150, and preferably 225. This suggestion was related to estimates of the local variograms. It is much more difficult to make such a statement about estimates of the theoretical variogram that applies to the population of realizations rather than to one particular study area. The uncertainties not only depend on the number of samples but also on the sampling scale. The uncertainties in the variogram parameters are almost independent of the sample size when the extent is small, while they strongly depend on the sample size for large extents. The uncertainty of the parameters also differs vastly with the number of variogram parameters to be estimated. One-hundred samples give a small uncertainty for the correlation length when it is the only parameter to be fitted and the extent is small. One thousand twenty-four samples can be too few for properly estimating a three-parameter variogram model if the extent is large. The exact range of relatively unbiased and certain es-
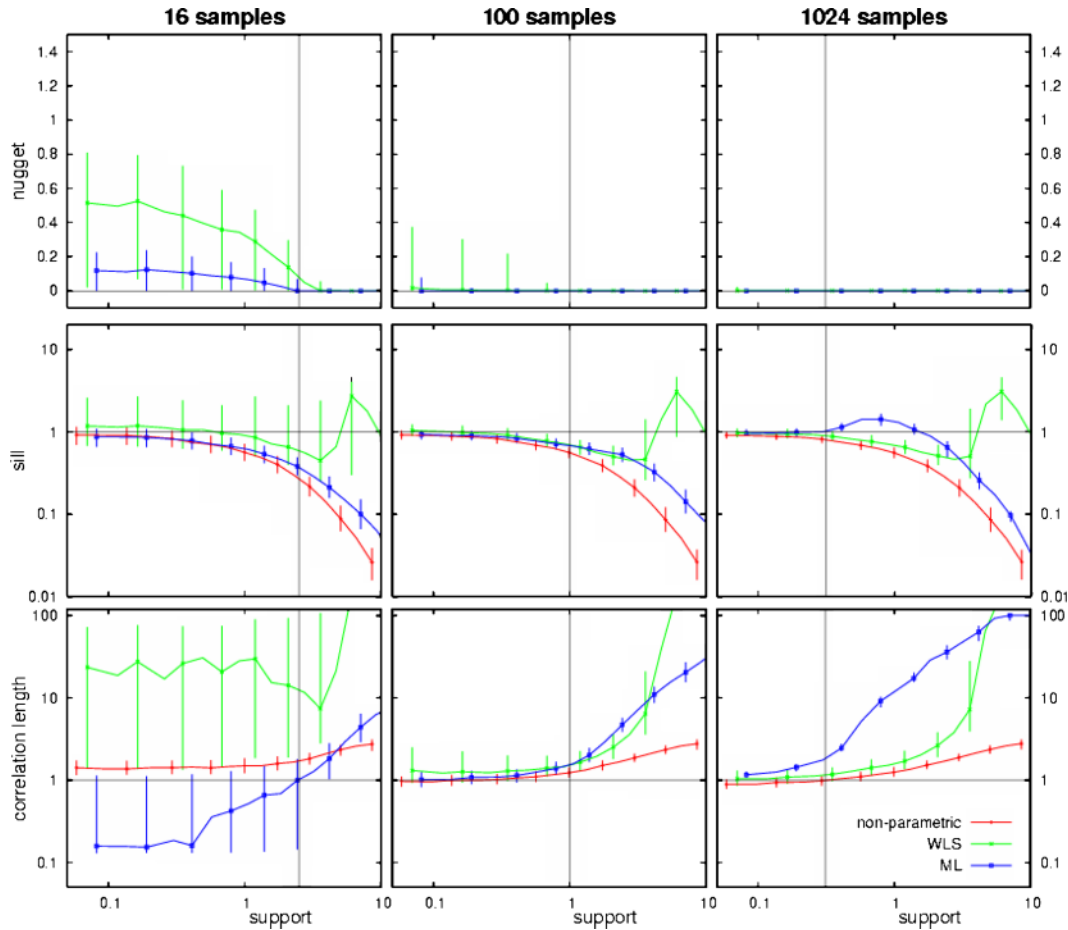
Fig. 11. (Top) Effect of support L_A^* on the estimated nugget $\hat{c}_0^*$, by the weighted least squares (WLS, green) and maximum likelihood (ML, blue) methods. (Middle) Effect of support L_A^* on the estimated sill $\hat{c}_s^*$ (WLS green, ML blue) and the sample variance s^{2*} (red). (Bottom) Effect of support L_A^* on the estimated correlation length $\hat{\lambda}^*$ (WLS green, ML blue) and the integral scale $\hat{J}^*$ (red). Gridded sampling (median with 25 and 75% quantiles as error bars). All are for the three-parameter model. $L_E^* = 10$, spacing L_S^* according to Eq. [8].

timates depends on the estimation method, the number of samples, and the spatial sampling arrangement. This range is the window where the sampling scales are commensurate with the scale of the underlying variability that is likely to be used in a sampling setup. A large number of samples both increases the range of unbiased estimates and reduces the uncertainty within this range. The parametric estimates are generally unbiased for nondimensional extents $L_E^* > 2$ and nondimensional spacings $L_S^* < 3$. Unbiasedness can also be achieved for larger spacings if the samples are taken at random locations. For a nested sampling arrangement one would also expect a wide range of unbiased estimates.

To make an a priori assessment of the suitability of a planned sampling setup, knowledge of the true correlation length or integral scale is required, as the analyses as conducted in this paper are based on sampling scales normalized by the true correlation length of the underlying process. There are a number of options to obtain such estimates. The first is an exploratory sampling along a transect or in a nested sampling arrangement to obtain initial estimates of the correlation length. The second is to use values from the literature of similar studies. Third, process reasoning can be used to obtain an educated

guess of the correlation lengths of the variables of interest in landscape soil processes. Seyfried and Wilcox (1995) proposed a framework of inferring correlation lengths from geological and physicals features that are important for hydrological modeling at different scales. Visual examination of outcrops and other soil-related features in the landscape can be a valuable guide for obtaining an order of magnitude of the scale at which soil processes operate in a particular environment.

It is in some cases possible to correct for the bias effects of the scale triplet. If the spacing is small, the extent is large, and the support is small, then the biases will be small. As the spacing gets larger, the extent smaller, and the support larger, the biases will increase. If the components of the scale triplet do not differ too much from the unbiased case, estimates of the biases of the variance and the correlation length can be back-calculated from the results of the simulations study in this paper. This is illustrated in Western and Blöschl (1999) and Skøien and Blöschl (2005). The uncertainty of these back-calculation methods increases with increasing spacing, decreasing extent, and increasing support. For the case of large supports it has also been suggested that one numerically back-calculates the point variogram by regulariza-
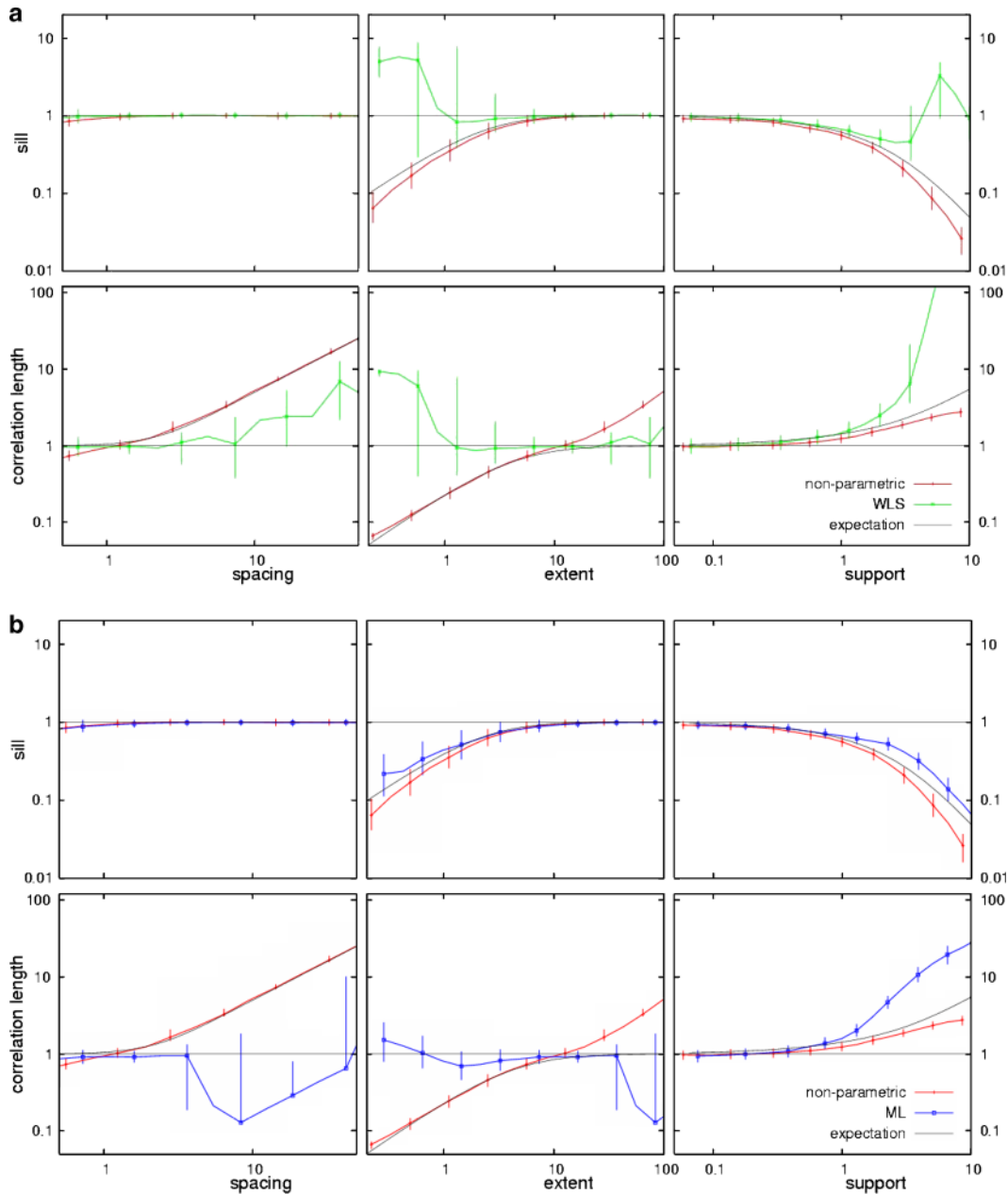
Fig. 12. Estimated sill and correlation length from (a) the weighted least squares (WLS) method (green) and (b) the maximum likelihood (ML) method (blue), together with variance and integral scale (red). The thin black lines correspond to the analytical expectations. All are for the two-parameter model, with 100 gridded samples.

tion methods, which can deal with more general cases than addressed here (Mockus, 1998; Skøien et al., 2003; Kyriakidis, 2004). In this analysis it has been assumed that the variable of interest aggregates linearly; that is, the value over a support area is simply the arithmetic average of the point values within this support area. Some of the variables of interest in soil hydrology do indeed aggregate linearly, such as soil moisture and soil thickness. Other variables, such as hydraulic conductivities, do not. More complicated aggregation rules exist for these parameters. These are dealt with in the abundant literature on upscaling and effective parameters (e.g., Blöschl and Sivapalan, 1995; Hopmans and Schoups, 2005). The support effects of this paper will in

this case only be an approximation of the real support effects.

The different estimation methods examined here have different characteristics. The non-parametric estimates are in most cases associated with the lowest uncertainties. They are on the other hand highly biased. The WLS method provides visually satisfying fits to the sample variograms. The uncertainty, however, tends to be large and increases with the number of variogram parameters to be estimated. The WLS method performs slightly better than the ML method for 16 samples. The ML method is computationally more demanding for large sample sizes because, for n data an $n \times n$ matrix has to be inverted for each iteration. The ML method gives better results

for randomly distributed samples with large spacings, but does not seem to be significantly better than the WLS method for other sampling scale combinations. In addition, estimated variograms tend to be too flat for small extents. Our results are consistent with analyses performed by Zimmerman and Zimmerman (1991, p.90) who noted that "'krigers' will sacrifice little by using the easy-to-compute OLS [Ordinary Least Squares] or … WLS … estimators, rather than the REML [Restricted Maximum Likelihood] estimators." Lark (2000) also compared ML estimates with estimates from fitting a theoretical variogram to the sample variogram. He concluded that ML estimates were often advantageous for short extents but this was not the case for large extents and large nuggets. For a certain experimental setup, he showed that an ML estimate from 60 samples provided an equally good fit to the underlying variogram as a theoretical model fitted to the sample variogram using 90–120 samples. Lark (2000) also concluded that a potential advantage of the ML method is lost when the data are contaminated with outliers, as most real data sets are.

Different opinions exist in the literature on whether variograms should be estimated automatically or hand-fitted. Webster and Oliver (2001) argued for automatic calibration, but suggested that the estimated variogram should be visually examined as well. However, a number of authors have argued in favor of visually fitting (AI-GEOSTATS, 2004). We used automatic fitting methods here because of the large number of variograms to be analyzed, and to allow a more objective interpretation of the results. On the other hand, the more complex the correlation structure, the more difficult it will be to visually fit a variogram. Visually fitting a variogram with few parameters such as that of Eq. [4] in one dimension is a simple task, but is more complicated if nested variograms and more dimensions are analyzed.

For small extents, typically where $L_E^* < 1$, strong trends will appear in the data. One could argue that for these cases a trend should be modeled explicitly and the geostatistical analysis should be performed on the residuals. There are several reasons why we have chosen to estimate the variogram directly from the data. First, we were interested in the population characteristics, and apparent trends are indeed a feature of the population characteristics for short extents. Second, there is no objective way of ascertaining whether there is a trend in the data or the strength of it (Leuangthong and Deutsch, 2004). Third, it would raise a range of questions beyond the purpose of this paper. Fourth, although in this paper it was clear that the variable was nonstationary within the extent of the data when extents were short, this is not always the case for real data. This point has been nicely illustrated by Gelhar (1993, p. 295), who presented the same variogram plotted on linear-linear and log-log scales. The former looks stationary, while the latter indicates that the variable is nonstationary. Di Federico and Neuman (1997) and Cintoli et al. (2005) suggested that the concept of truncated power variograms can be used to explain these sampling effects.

In vadose zone applications at the landscape scale, there are a range of variables that are of interest, includ-ing soil water content, concentrations of chemical or biological tracers, as well as hydraulic soil characteristics. In examining these variables, it is a key concept that sampling always involves some sort of filtering (Cushman, 1984, 1987). It is important to take these filtering aspects into account when interpreting both observations and model estimates of these variables. The major implication of this study for soil hydrology is that sampling scales will likely have an important effect on the variogram parameters estimated from vadose zone data at the landscape scale. Sampling scales will introduce biases and random errors if these scales depart from the scales of the underlying process. Their magnitude is controlled by the sampling scale triplet relative to the scale of the underlying process.

REFERENCES

AI-GEOSTATS. 2004. Discussion on "Automated variogram modelling" [Online message board]. Messages 1524–1550. Available at http://groups.yahoo.com/group/new_ai_geostats/messages (accessed 5 Sept. 2005, verified 20 Dec. 2005).

Beckie, R. 1996. Measurement scale, network sampling scale, and groundwater model parameters. Water Resour. Res. 32:65–76.

Blöschl, G. 1999. Scaling issues in snow hydrology. Hydrol. Processes 13:2149–2175.

Blöschl, G., and M. Sivapalan. 1995. Scale issues in hydrological modelling—A review. Hydrol. Processes 9:251–290.

Cintoli, S., S.P. Neuman, and V. Di Federico. 2005. Generating and scaling fractional Brownian motion on finite domains. Geophys. Res. Lett. 32:L08404. doi:10.1029/2005GL022608.

Cressie, N. 1985. Fitting variogram models by weighted least squares. Math. Geol. 17:563–586.

Cressie, N. 1991. Statistics for spatial data. John Wiley and Sons, New York.

Cressie, N., and D. Hawkins. 1980. Robust estimation of the variogram. I. J. Int. Assoc. Math. Geol. 12:115–125.

Cushman, J.H. 1984. On unifying the concepts of scale, instrumentation, and stochastics in the development of multiphase transport theory. Water Resour. Res. 20:1668–1676.

Cushman, J.H. 1987. More on stochastic models. Water Resour. Res. 23:750–752.

Di Federico, V., and S.P. Neuman. 1997. Scaling of random fields by means of truncated power variograms and associated spectra. Water Resour. Res. 33:1075–1085.

Duan, Q., S. Sorooshian, and V.K. Gupta. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28:1015–1031.

Dungan, J.L., J.N. Perry, M.R.T. Dale, P. Legendre, S. Citron-Pousty, M.-J. Fortin, A. Jakomulska, M. Miriti, and M.S. Rosenberg. 2002. A balanced view of scale in spatial statistical analysis. Ecography 25:626–640.

Gelhar, L.W. 1993. Stochastic subsurface hydrology. Prentice Hall, Englewood Cliffs, NJ.

Gorsich, D.J., and M.G. Genton. 2000. Variogram model selection via nonparametric derivative estimation. Math. Geol. 32:249–270.

Haws, N.W., B. Liu, C.W. Boast, P.S.C. Rao, E.J. Kladivko, and D.P. Franzmeier. 2004. Spatial variability and measurement scale of infiltration rate on an agricultural landscape. Soil Sci. Soc. Am. J. 68:1818–1826.

Hopmans, J.W., and G.H. Schoups. 2005. Soil water flow at different spatial scales. p. 999–1010. In G. Anderson (ed.) Encyclopedia of hydrological sciences. Vol. M. Wiley, Chichester, UK.

Journel, A.G., and C.J. Huijbregts. 1978. Mining geostatistics. Academic Press, London.

Kitanidis, P.K. 1983. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. Water Resour. Res. 19:909–921.

Kitanidis, P.K., and R.W. Lane. 1985. Maximum likelihood parameter estimation of hydrologic spatial processes by the Gauss-Newton method. J. Hydrol. (Amsterdam) 79:53–71.

Kyriakidis, P.C. 2004. A geostatistical framework for area-to-point spatial interpolation. Geograph. Anal. 36:259–289.

Lark, R.M. 2000. Estimating variograms of soil properties by the method-of-moments and maximum likelihood. Eur. J. Soil Sci. 51: 717–728.

Leuangthong, O., and C.V. Deutsch. 2004. Transformation of residuals to avoid artifacts in geostatistical modelling with a trend. Math. Geol. 36:287–305.

Mantoglou, A., and J.L. Wilson. 1981. Simulation of random fields with the turning bands method. Dep. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Mantoglou, A., and J.L. Wilson. 1982. The turning bands method for simulation of random fields using line generation by a spectral method. Water Resour. Res. 18:1379–1394.

Matheron, G. 1965. Les variables regionalisées et leur estimation. Masson, Paris.

McBratney, A.B. 1992. On variation, uncertainty and informatics in environmental soil management. Aust. J. Soil Res. 30:913–935.

Menke, W. 1989. Geophysical data analysis: Discrete inverse theory. Academic Press, New York.

Mockus, A. 1998. Estimating dependencies from spatial averages. J. Comput. Graph. Stat. 7:501–513.

Mohanty, B.P., J.S. Famiglietti, and T.H. Skaggs. 2000. Evolution of soil moisture spatial structure in a mixed vegetation pixel during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resour. Res. 36:3675–3686.

Nyquist, H. 1924. Certain factors affecting telegraph speed. Bell Syst. Tech. J. 3:324–346.

Oliver, M.A., and A.L. Khayrat. 2001. A geostatistical investigation of the spatial variation of radon in soil. Comput. Geosci. 27: 939–957.

Omre, H. 1984. The variogram and its estimation. p. 107–125. *In* G. Verly et al (ed.) Geostatistics for natural resources characterization. Vol. 1. D. Reidel Publishing Company, Dordrecht, The Netherlands.

Ortiz, C.J., and C.V. Deutsch. 2002. Calculation of uncertainty in the variogram. Math. Geol. 34:169–183.

Pardo-Igúzquiza, E. 1998. Maximum likelihood estimation of spatial covariance parameters. Math. Geol. 30:95–108.

Pardo-Igúzquiza, E., and P. Dowd. 2001. Variance-covariance matrix of the experimental variogram: Assessing variogram uncertainty. Math. Geol. 33:397–419.

R Development Core Team. 2004. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ribeiro, P.J., Jr., and P.J. Diggle. 2001. geoR: A package for geostatistical analysis. R-NEWS 1:15–18.

Romshoo, S.A. 2004. Geostatistical analysis of soil moisture measurements and remotely sensed data at different spatial scales. Environ. Geol. 45:339–349.

Russo, D., and W.A. Jury. 1987. A theoretical study of the estimation of the correlation scale in spatially variable fields. 1. Stationary fields. Water Resour. Res. 23:1257–1268.

Seyfried, M.S., and B.P. Wilcox. 1995. Scale and the nature of spatial variability: Field examples having implications for hydrological modeling. Water Resour. Res. 31:173–184.

Shafer, J.M., and M.D. Varljen. 1990. Approximation of confidence limits on sample semivariograms from single realisations of spatially correlated random fields. Water Resour. Res. 26:1787–1802.

Skøien, J.O., and G. Blöschl. 2006. Sampling scale effects in two dimensional random fields and implications for environmental modelling. Environ. Monit. Assess. (In press).

Skøien, J.O., G. Blöschl, and A.W. Western. 2003. Characteristic space-time scales in hydrology. Water Resour. Res. 39:1304. doi:10.1029/2002WR001736.

Taylor, G.I. 1921. Diffusion by continuous movements. Proc. London Math. Soc. 20:196–211.

Webster, R. 2000. Is soil variation random? Geoderma 97:149–163.

Webster, R., and M.A. Oliver. 1992. Sample adequately to estimate variograms of soil properties. J. Soil Sci. 43:177–192.

Webster, R., and M.A. Oliver. 2001. Geostatistics for environmental scientists. Wiley, Chichester, UK.

Western, A.W., and G. Blöschl. 1999. On the spatial scaling of soil moisture. J. Hydrol. (Amsterdam) 217:203–224.

Western, A.W., S.-L. Zhou, R.B. Grayson, T.A. McMahon, G. Blöschl, and D.J. Wilson. 2004. Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. J. Hydrol. (Amsterdam) 286:113–134.

Woodbury, A.D., and E.A. Sudicky. 1991. The geostatistical characteristics of the Borden Aquifer. Water Resour. Res. 27:533–546.

Zimmerman, D.L., and M.B. Zimmerman. 1991. A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. Technometrics 33:77–91.