Quality measures for soil surveys by lognormal kriging

R.M. Lark[a*], D.J. Lapworth[b].

[a]*British Geological Survey, Keyworth, Nottinghamshire NG12 5GG, U.K.,* [b]*British Geological Survey, Maclean Building, Wallingford, Oxfordshire, OX10 8BB, U.K.*

Abstract

If we know the variogram of a random variable then we can compute the prediction error variances (kriging variances) for kriged estimates of the variable at unsampled sites from sampling grids of different design and density. In this way the kriging variance is a useful pre-survey measure of the quality of statistical predictions, which can be used to design sampling schemes to achieve target quality requirements at minimal cost. However, many soil properties are lognormally distributed, and must be transformed to logarithms before geostatistical analysis. The predicted values on the log scale are then back-transformed. It is possible to compute the prediction error variance for a prediction by this lognormal kriging procedure. However, it does not depend only on the variogram of the variable and the sampling configuration, but also on the conditional mean of the prediction. We therefore cannot use the kriging variance directly as a pre-survey measure of quality for geostatistical surveys of lognormal variables. In this paper we present an alternative. First we show how the limits of a prediction interval for a variable predicted by lognormal kriging can be expressed as dimensionless quantities, proportions of the unknown median of the conditional distribution. This scaled prediction interval can be used as a presurvey quality measure since it depends only on the sampling configuration and the variogram of the log-transformed variable. Second, we show how a similar scaled prediction interval can be computed for the median value of a lognormal variable across a block, in the case of block kriging. This approach is then illustrated using variograms of lognormally distributed data on concentration of elements in the soils of a part of eastern England.

Keywords: lognormal kriging, prediction interval, sampling, geostatistics.

---
*Corresponding author: *E-mail address*: mlark@nerc.ac.uk (R.M. Lark).

1. Introduction

There is a growing awareness of the need to manage the soil sustainably, and as a result regulatory frameworks have been developed to ensure that soil quality is maintained (e.g. European Commission, 2006). As Bone et al. (2010) observe, the assessment of soil quality is challenging, potentially costly and prone to uncertainty because of the variability of soil material. It is therefore important that sampling schemes for soil assessment are carefully designed. De Gruijter et al. (2006) discuss how sampling can be planned so that questions about the soil are answered satisfactorily and efficiently. It is necessary to make best use of costly field and laboratory effort, and the results from sampling and analysis must be sufficiently precise to meet the end-user's requirements. De Gruijter et al (2006) emphasize the importance of clearly identifying what these requirements are before the survey is planned. For example, the *target quantity* that we want to know might be the mean value of some variable across a region of interest, and an estimate of this is usually best achieved by an appropriately designed probability sample which entails randomization. If, alternatively, the user wants a set of local predictions (perhaps presented as a contour map) then this requires a more or less regular array of sample locations, and appropriate model-based statistical analyses. Having identified the nature of the question that sampling is to answer, we must also have some idea of how reliable the answer must be. This can be expressed by what de Gruijter et al. (2006) call *quality measures*.

A quality measure is a measure of the precision of an estimate from sample data. Once we have some data we can compute estimates of target quantities from them (e.g. means), and associated quality measures (e.g. confidence intervals). These are post-survey quality measures, which tell us, and users of the information, how well we have done. What we require for planning sampling are pre-survey quality measures, which tell us how well we can expect to do given a certain survey effort. Usually we can only approximate pre-survey quality measures (they may depend on estimates of

values such as the variance of the target quantity in the population of interest that we can only approximate before sampling). Such pre-survey quality measures may be the expected width of the confidence interval for a target quantity, or the statistical power with which we can detect a change in the soil (de Gruijter et al, 2006; Lark, 2009; Brus and Noij, 2008).

Ideally we identify a quality measure that is appropriate for a particular sampling problem, and which can be approximated, pre-survey, from available information. We also ask the data user to specify values of the quality measure that are acceptable for their purposes. It should then be possible to plan a sampling campaign that will return information of suitable quality at acceptable minimal cost, or to show the user that this is not possible, and that it is necessary either to increase the budget to permit the collection of more samples or to accept that less precise estimates will be possible than originally hoped.

The kriging methods introduced to soil science by Burgess and Webster (1980), and further developments of these, are routinely used to produce local predictions of soil properties when such predictions are the required outcome from a soil inventory. Geostatistical methods are model-based in that they invoke an underlying random variable that is held to be realized in observed data, rather than depending on randomized sampling. The spatial dependence of this random variable is modelled by the variogram function. Local predictions are obtained as weighted averages of neighbouring observations of the variable, the weights being selected to minimize the expected squared error of the predictions. This quantity, called the kriging variance, is reported along with the prediction. It is a useful quality measure. Note that local predictions by point kriging are made on the original quasi-point support of the data — our observations are made on soil cores or similar specimens that are of very small dimensions by comparison to the region under study. As an alternative to point kriging we may estimate the mean value of the target variable over some region or block, which may be

a regular rectilinear panel or an irregular region such as a field or similar management unit. This is called block kriging.

Geostatistical prediction by kriging is based on a random statistical model of the variable of interest which is inferred from data. Our data are treated as a realization of the underlying random model (de Gruijter et al., 2006). In kriging our target quantity is a point or block value specific to the realization, the block value is the spatial mean of the variable over the block's extent. The measures of uncertainty (kriging variances) are derived over the model distribution conditional on the observations.

Given the variogram function, the kriging variance (point or block kriging) for some variable at a particular location depends only on the configuration of sample sites. This makes the kriging variance a useful pre-survey quality measure. If we have an estimate of the variogram, perhaps from a survey of a neighbouring region, we can identify a sample network which ensures that the kriging variances of local predictions fall within an acceptable range. This was demonstrated by McBratney et al (1981), and their approach has been applied to the design of soil surveys (Di et al., 1989; van Groenigen et al., 1999) including cases where the kriging prediction includes an external drift modelled by covariates such as remote sensor data (Brus and Heuvelink, 2007). More recently this work has been developed for the optimization of spatial surveys including both variogram estimation and prediction by kriging (Marchant and Lark, 2006; Zhu and Stein, 2006; Marchant and Lark, 2007). Note that sound inference from the variogram requires that it has been estimated reliably. When data are prone to including outliers then robust variogram estimators may be needed, and the resulting model must be validated (Lark, 2000).

It is commonly found that soil and other geochemical variables do not appear to be normally distributed (White et al., 1987; Allègre and Lewin, 1995). This is best judged by exploratory statistics, such as the coefficient of skewness, and histograms of the data (Webster and Oliver, 2007). In these circumstances the data should be

transformed to a scale of measurement on which an underlying normally distributed random variable can plausibly be assumed. Geostatistical predictions can be obtained on this new scale and then back-transformed to the scale of measurement. This is called trans-Gaussian kriging (Cressie, 1993). A common case is lognormal kriging, when the data are transformed to logarithms. Let Y be the normal variable obtained by transformation of our original variable, Z, to natural logarithms. The ordinary point kriging of Y at location $\mathbf{x}_0$ is the conditional mean of the variable $Y(\mathbf{x}_0)$, conditional on the observed values used for prediction, the random model (variogram) the assumption of a fixed but unknown local mean of Y and the assumption that Y is a normal random variable (Stein, 1999). The conditional distribution of $Y(\mathbf{x}_0)$ has variance $\sigma_{\mathrm{K}}^2(\mathbf{x}_0)$, the kriging variance. However, for scientific or practical purposes we generally require predictions on the original scale. The ordinary point kriging prediction of Y is back-transformed to Z, the corresponding variable on the original scale of measurement by

$$\tilde{Z}(\mathbf{x}_0) \ = \exp\left\{\tilde{Y}(\mathbf{x}_0) + \frac{\sigma_{\mathrm{K}}^2(\mathbf{x}_0)}{2} - \psi(\mathbf{x}_0)\right\}, \tag{1}$$

where $\psi(\mathbf{x}_0)$ is a Lagrange multiplier obtained in the solution of the kriging equations and $\tilde{Z}$ and $\tilde{Y}$ denote the kriging predictions of the respective random variables. The prediction error variance on the original scale of measurement can be written as

$$\left[\exp\left\{2\mu_Y + \sigma_{\mathrm{K}}^2(\mathbf{x}_0)\right\}\right] \times$$
$$\left[\exp\left\{\sigma_{\mathrm{K}}^2(\mathbf{x}_0)\right\} + \exp\left\{\mathrm{Var}\left(\tilde{Y}(\mathbf{x}_0)\right)\right\} - 2\exp\left\{\mathrm{Cov}\left(Y(\mathbf{x}_0), \tilde{Y}(\mathbf{x}_0)\right)\right\}\right], \tag{2}$$

where μ_Y is the mean of Y and $\mathrm{Var}(\cdot)$ and $\mathrm{Cov}(\cdot, \cdot)$ denote, respectively the variance and covariance of the terms in brackets (Cressie, 1993). The key property of this latter expression is that the variance of the prediction depends on the mean of the variable. For this reason, unlike ordinary point or block kriging on the untransformed data, we cannot express the kriging variance as a quality measure dependent only on the variogram and the sampling design. In the log-normal case the kriging variance is

therefore only useful as a post-survey quality measure, and cannot be used to select among different sampling designs before we have sampled a particular region.

The aim of this paper is to explore and demonstrate alternative quality measures that could be used for pre-survey planning of sampling for lognormally distributed variables. Some approaches are proposed for ordinary point and block kriging, and then illustrated with soil data from a baseline geochemical survey of part of eastern England.

2. Theory

2.1. The proposed quality measures: standardized prediction intervals and quantiles

As seen above, the kriging variance of a normally distributed random variable is a useful pre-survey quality measure for a sampling scheme because it depends only on the variogram of the variable and the sampling configuration. By contrast the mean square prediction error (kriging variance) at location $\mathbf{x}_0$ on the untransformed scale depends, *inter alia* on $\mathrm{E}[Z(\mathbf{x}_0)]$ conditional on the observations, so will not serve as a pre-survey quality measure. In this paper we propose quality measures based on prediction intervals rather than variances. A prediction interval of some random quantity X, $(L_\alpha(X), U_\alpha(X))$, is an interval with an assigned probability α such that

$$\mathrm{Prob}\left[L_\alpha(X) < X < U_\alpha(X)\right] = 1 - \alpha. \tag{3}$$

In this paper we consider prediction intervals which are symmetric in the sense that

$$\mathrm{Prob}\left[L_\alpha(X) < X < \mathrm{median}(X)\right] = \mathrm{Prob}\left[\mathrm{median}(X) < X < U_\alpha(X)\right] = \frac{1 - \alpha}{2}, \tag{4}$$

where $\mathrm{median}(\cdot)$ denotes the median of a random variable. We show how we can compute the upper and lower bounds of prediction intervals for target quantities that might be obtained by lognormal point or block kriging. In the case of point kriging the target quantity is the unobserved value $Z(\mathbf{x}_0)$. In the case of block kriging the target quantity is the median value of the variable Z over a particular block. We then show

how these bounds can be used as pre-survey quantity measures by expressing them as proportions, in the point kriging case of the conditional median value of $Z(\mathbf{x}_0)$, in the block kriging case of the conditional median value of the variable Z across the block.

2.2. Point kriging

We are used to the mean and variance as summary statistics for normal random variables, but they are rather less useful for log-normally distributed variables since they are dominated by values from the upper tail of the distribution. This is why the geometric mean or the median of a log-normal random variable is commonly preferred as a summary statistic, and the conditional median of a log-normal random variable has been proposed as a more useful target quantity for spatial prediction than the conditional mean (Pawlowsky-Glahn and Olea, 2004; Tolosana-Delgado and Pawlowsky-Glahn, 2007). Now because exponentiation, the back-transform from a value of Y to a value of Z is a strictly non-decreasing function, the exponentiation of the conditional median of $Y(\mathbf{x}_0)$ gives the conditional median of $Z(\mathbf{x}_0)$. Since the mean and median of a normal random variable are identical, the simple back-transform of $\tilde{Y}(\mathbf{x}_0)$ by exponentiation gives a median-unbiased estimate of a log-normal variable $Z(\mathbf{x}_0)$ (Chilès and Delfiner, 1999).

As Chilès and Delfiner (1999) point out, this result for back transformation holds for any percentile of $Y(\mathbf{x}_0)$. As a result, we can back-transform the end-members of a prediction interval on the transformed scale, $L_\alpha(Y), U_\alpha(Y)$ to find corresponding end-members of a prediction interval, with the same value of α, on the original scale of measurement. If, for example, we are interested in a 90% prediction interval ($\alpha = 0.1$), then the end-members of such an interval, symmetric about the conditional median of Z in the sense of Eq. (4), are $\exp\left\{\tilde{Y}(\mathbf{x}_0) - 1.64\sigma_\mathrm{K}(\mathbf{x}_0)\right\}$ and $\exp\left\{\tilde{Y}(\mathbf{x}_0) + 1.64\sigma_\mathrm{K}(\mathbf{x}_0)\right\}$.

Now, it is clear that the prediction interval

$$\left(\exp\left\{\tilde{Y}(\mathbf{x}_0) - 1.64\sigma_\mathrm{K}(\mathbf{x}_0)\right\}, \exp\left\{\tilde{Y}(\mathbf{x}_0) + 1.64\sigma_\mathrm{K}(\mathbf{x}_0)\right\}\right)$$

can be rewritten as

$$\left(\exp\left\{\tilde{Y}(\mathbf{x}_0)\right\}\exp\left\{-1.64\sigma_K(\mathbf{x}_0)\right\},\exp\left\{\tilde{Y}(\mathbf{x}_0)\right\}\exp\left\{+1.64\sigma_K(\mathbf{x}_0)\right\}\right)$$

We can therefore express the limits of the prediction interval as *standardized limits*, $L_{s,0.1}$ and $U_{s,0.1}$ which are dimensionless values, proportions of the conditional median of Z:

$$L_{s,0.1} = \exp\left\{-1.64\sigma_K(\mathbf{x}_0)\right\}$$

$$U_{s,0.1} = \exp\left\{1.64\sigma_K(\mathbf{x}_0)\right\}. \tag{5}$$

Of course we could compute standardized limits for any value of α that seems appropriate. In this paper we use $\alpha = 0.1$ throughout.

In summary, the target quantity for the geostatistical survey that we consider here is the conditional median of the variable of interest at an unsampled site, and our quality measure is defined in terms of the prediction interval of this quantity over the model distribution. We propose that the standardized limits in Eq. (5) are used to derive pre-survey quality measures for lognormal point kriging, since they depend only on the kriging variance of the transformed variable, and so only on its variogram and the distribution of sample points. The lower limit, $L_{s,0.1}$ could be a useful quality measure in itself. The possible values of the lower limit are constrained, $L_{s,0.1} \in (0, 1)$ and it should be intuitively clear to the user that the closer it is to 1 the better the quality of the prediction. One might specify, for example, that a target value for $L_{s,0.1}$ is 0.75, i.e. the lower bound of the prediction interval is no less than 75% of the conditional median. By contrast $U_{s,0.1}$ has no upper bound, but we might specify some target maximum value, for example that the upper limit exceeds the conditional median by a proportion, $U_{s,0.1} - 1$, no larger than 0.75. Another possible criterion is the width of the standardized prediction interval

$$U_{s,0.1} - L_{s,0.1},$$

8

but the disadvantage of this is that it conceals the asymmetry of the interval about the conditional median. A simple plot of $U_{s,0.1}$ and $L_{s,0.1}$ against the sample density of a grid from which kriged estimates are notionally derived will present this information in a readily accessible way, showing how the scaled prediction interval shrinks in response to increased survey effort.

2.3. Block kriging

2.3.1. Lognormal block kriging. Block lognormal kriging, which entails a change of support from quasi-point observations (e.g. on soil cores) to larger regions for which we require predictions, as well as the non-linear transformation of data, is more challenging than point lognormal kriging. It has recently received some attention in the geostatistical literature (Cressie, 2006; Paul and Cressie, 2011). In particular Cressie (2006) draws attention to a proposal by Matheron (1974) that the block kriged estimate of the mean value of Z for block $\mathcal{B}$ be formed as the integral of unbiased point kriging estimates over $\mathcal{B}$:

$$\tilde{Z}^{\mathcal{B}} = \frac{1}{|\mathcal{B}|} \int_{\mathbf{x} \in \mathcal{B}} \exp\left\{\tilde{Y}(\mathbf{x}) + \frac{\sigma_{\mathrm{K}}^2(\mathbf{x})}{2} - \psi(\mathbf{x})\right\} \mathrm{d}\mathbf{x} , \tag{6}$$

where the integral is over the dimensions of the block, and $|\mathcal{B}|$ is the Lebesgue measure of $\mathcal{B}$. Cressie (2006) discusses how prediction error variances can be formed for these estimates. The procedure is computationally demanding, and, as with point kriging, the prediction error variances are not independent of the block mean so are not applicable as pre-survey quality measures in the sense of this paper.

2.3.2. Proposed quality measures based on the block median In this paper we present some pre-survey quality measures based on an estimate of the median value of a property across a block $\mathcal{B}$, obtained from the ordinary point kriging estimates of the transformed variable at locations that comprise a discrete approximation to $\mathcal{B}$:

$$\mathbf{X}^{\mathcal{B}} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\} , \ \mathbf{x}_i \in \mathcal{B} \ \forall \ i = 1, 2, \ldots, N.$$

The value of the transformed variable Y at $\mathbf{x}_i$, the ith location in the discrete approximation, is $Y_i^{\mathcal{B}}$. We treat the ordinary kriging prediction of it, $\tilde{Y}_i^{\mathcal{B}}$, as a random variable with the following structure:

$$\tilde{Y}_i^{\mathcal{B}} = \mu_Y^{\mathcal{B}} + \eta_{Y,i}^{\mathcal{B}}, \tag{7}$$

where $\mu_Y^{\mathcal{B}}$ is the mean value of Y across $\mathcal{B}$. Note that $\mu_Y^{\mathcal{B}}$ is the spatial mean across the block, and so it is a random quantity between realizations of the random variable, Y. The random variable $\eta_{Y,i}^{\mathcal{B}}$ is the deviation between the conditional expectation of $Y_i^{\mathcal{B}}$ and the the block spatial mean $\mu_Y^{\mathcal{B}}$.

Equation (7) describes a random variable, in practice we have one realization:

$$\tilde{y}_i^{\mathcal{B}} = m_Y^{\mathcal{B}} + e_{Y,i}^{\mathcal{B}}, \tag{8}$$

where $\tilde{y}_i^{\mathcal{B}}$ is the ordinary kriging prediction at the ith location, and $m_Y^{\mathcal{B}}$ is the (unknown) spatial mean across the block. Since the average of $\tilde{y}_i^{\mathcal{B}}$ over a sufficiently dense discrete approximation of the block tends the to conditional expectation of the block spatial mean, the mean of $e_{Y,i}^{\mathcal{B}}$ tends to the kriging error of the block spatial mean. In this paper we consider a predictor of the median value of our variable over the block, conditional on the ordinary kriging predictions, and hence derived from the predictions $\tilde{y}_i^{\mathcal{B}}$ over $\mathbf{X}^{\mathcal{B}}$. As with any kriging prediction the uncertainty is considered over the underlying random model, to which we now return.

We assume that the mean of $\eta_{Y,i}^{\mathcal{B}}$ is zero (the point kriging is unbiased) and its variance is denoted $\sigma_{\eta,i}^2$. It is not guaranteed that ordinary kriging is unbiased, but we know that the bias is minimized, subject to the constraints that allow ordinary kriging to handle an unknown local mean (Chilès and Delfiner, 1999). If, in addition, we assume normality of the within-block variation of $\eta_{Y,i}^{\mathcal{B}}$ then we can write:

$$\tilde{Y}_i^{\mathcal{B}} \sim \mathcal{N}\left\{\mu_Y^{\mathcal{B}}, \left(\sigma_{\eta,i}^2\right)\right\}. \tag{9}$$

In block kriging, following Matheron's proposal, we could back-transform each estimate $\tilde{Y}_i^{\mathcal{B}}$ without bias to an estimate of the conditional mean of the lognormal

variable at the ith location in the discrete approximation, $\tilde{Z}_i^{\mathcal{B}}$ and these are averaged over all locations in the discrete approximation to estimate the block conditional mean. Assume, however, that we form an alternative biased estimate, $\dot{Z}^{\mathcal{B}}$ thus

$$
\begin{aligned}
\dot{Z}^{\mathcal{B}} &= \frac{1}{N} \sum_{i=1}^{N} \exp\left\{\tilde{Y}_i^{\mathcal{B}}\right\} \\
&= \frac{1}{N} \sum_{i=1}^{N} \exp\left\{\mu_Y^{\mathcal{B}} + \eta_{Y,i}^{\mathcal{B}}\right\} \\
&= \exp\left\{\mu_Y^{\mathcal{B}}\right\} \frac{1}{N} \sum_{i=1}^{N} \exp\left\{\eta_{Y,i}^{\mathcal{B}}\right\}.
\end{aligned} \tag{10}
$$

We can see that the first term on the right-hand side of Eq. (10) is the median value of the lognormally distributed variable across the block given the assumption of normality of $\eta_{Y,i}^{\mathcal{B}}$. Because $\eta_{Y,i}^{\mathcal{B}}$ has mean zero, the expected value of $\dot{Z}^{\mathcal{B}}$, following from the familiar properties of the lognormal distribution, is

$$
\mathrm{E}\left[\dot{Z}^{\mathcal{B}}\right] = \exp\left\{\mu_Y^{\mathcal{B}}\right\} \frac{1}{N} \sum_{i=1}^{N} \exp\left\{\frac{\sigma_{\eta,i}^2}{2}\right\}. \tag{11}
$$

which provides us with an unbiased estimator of the block median for some particular realization, $\check{z}$:

$$
\widehat{\check{z}} = \frac{\sum_{i=1}^{N} \exp\left\{\tilde{y}_i^{\mathcal{B}}\right\}}{\sum_{i=1}^{N} \exp\left\{\frac{\sigma_{\eta,i}^2}{2}\right\}}. \tag{12}
$$

Note that $\widehat{\check{z}}$ is an unbiased estimate of the mean value of the block median for our particular realization over the model distribution, depending on the variances $\sigma_{\eta,i}^2$, $i = 1, 2, \ldots N$. If we were to generate $\mathbf{e}^{\mathcal{B}} = \left\{e_{Y,1}^{\mathcal{B}} \ldots, e_{Y,N}^{\mathcal{B}}\right\}$, a realization of the random variate $\boldsymbol{\eta}^{\mathcal{B}}$ we could then compute

$$
s = \frac{\sum_{i=1}^{N} \exp\left\{e_{Y,i}^{\mathcal{B}}\right\}}{\sum_{i=1}^{N} \exp\left\{\frac{\sigma_{\eta,i}^2}{2}\right\}}, \tag{13}
$$

which is a realization from the model distribution of $\widehat{\check{Z}}$ scaled to dimensionless values that are proportions of the unknown block median. By generating multiple such realizations, and computing appropriate percentiles of their distribution, we can obtain

approximate prediction intervals expressed, as in the point kriging case, as proportions of the unknown target quantity. This makes these intervals suitable pre-survey quality measures. In order to do this we need a way to generate realizations of $\boldsymbol{\eta}^{\mathcal{B}}$, and obtain the unknown variances $\sigma_{\eta,i}^2$, $i = 1, 2, \ldots N$. This is described in the next section.

In summary, for the block kriging case our quantity of interest is the spatial median of variable Z across the block, the median over the distribution of realized values, and Eq. (12) allows us to compute its expectation, conditional on the data, over the distribution of the random model. As quality measures we propose standardized prediction intervals of this expectation over the random distribution.

2.3.3. Implementation For some particular block, $\mathcal{B}$, we have a discrete approximation $\mathbf{X}^{\mathcal{B}}$. At any location in this approximation, $\mathbf{x}_i$ the conditional expectation $\tilde{Y}_i^{\mathcal{B}}$ is estimated by ordinary kriging. All these predictions are obtained by kriging from the observations at a common set of M locations, $\mathbf{X}^{\mathcal{P}} = \mathbf{x}_{\mathcal{P},1}, \ldots, \mathbf{x}_{\mathcal{P},M}$.

The most straightforward way to generate a realization of $\boldsymbol{\eta}^{\mathcal{B}}$ is by a numerical approximation. We generate a realization of the random variable Y by a standard simulation method at a set of locations that comprises the prediction set, $\mathbf{X}^{\mathcal{P}}$ and a random set of ν locations drawn at random from across the block: $\mathbf{X}^{\mathcal{BR}}$. We may estimate $m_y^{\mathcal{B}}$, the spatial mean across the block for the particular realization, by the arithmetic average of the simulated values at locations in $\mathbf{X}^{\mathcal{BR}}$. We can also derive kriged estimates at points in $\mathbf{X}^{\mathcal{B}}$, $\tilde{y}_i^{\mathcal{B}}$, $i = 1, \ldots, N$, by ordinary kriging from the simulated values at locations in $\mathbf{X}^{\mathcal{P}}$. These values are then substituted into Eq (8) to provide a realization $\mathbf{e}^{\mathcal{B}} = \left\{ e_{Y,1}^{\mathcal{B}} \ldots, e_{Y,N}^{\mathcal{B}} \right\}$. Because we assume that the expectation of $\eta_i^{\mathcal{B}}$ is zero, we can obtain estimates of $\sigma_{\eta,i}^2$ by calculating the mean square value of many such realizations. Because we are substituting an estimate of $m_y^{\mathcal{B}}$ for the unknown value for each realization there will be some bias, but this is reduced by using a large number, ν, of values within the block to obtain the estimate, not including the points in $\mathbf{X}^{\mathcal{B}}$.

Our procedure to compute scaled limits for the prediction interval of a block median $L^{\mathcal{B}}_{s,\alpha}$, $U^{\mathcal{B}}_{s,\alpha}$ is therefore as follows.

1. Generate the coordinates of the N points in the discrete approximation to the block, $\mathbf{X}^{\mathcal{B}}$.

2. Generate the coordinates of ν points in the block in the set $\mathbf{X}^{\mathcal{BR}}$. These are obtained by probability sampling with a uniform inclusion probability density across the block.

3. Compute the ordinary kriging weights for predicting Y at all locations in $\mathbf{X}^{\mathcal{B}}$ from the locations in $\mathbf{X}^{\mathcal{BR}}$.

4. Compute the $(M+\nu) \times (M+\nu)$ covariance matrix for Y between all $M+\nu$ locations in the union of sets $\mathbf{X}^{\mathcal{P}}$ and $\mathbf{X}^{\mathcal{BR}}$. Generate a realization of Y at these locations by premultiplying a $(N+\nu) \times 1$ standard normal variate by the lower-triangular Cholesky factor of the covariance matrix (e.g. Webster and Oliver, 2007).

5. Estimate the block mean by the average of the simulated values at the ν locations in $\mathbf{X}^{\mathcal{BR}}$.

6. Calculate the ordinary kriging estimates at the locations in $\mathbf{X}^{\mathcal{B}}$ by applying the kriging weights computed at step (3) to the simulated values at the M locations in $\mathbf{X}^{\mathcal{P}}$.

7. Substitute the values obtained at steps (5) and (6) into to Eq. (8) to obtain the elements of the realization $\mathbf{e}^{\mathcal{B}}$. Store this realization and then reiterate steps (2) to (7) a large number, κ, of times.

8. After computing all iterations, compute estimates of the variances $\sigma^2_{\eta,i}$, $i = 1, \ldots, N$ by the mean square value of the corresponding values of $\mathbf{e}^{\mathcal{B}} = \left\{ e^{\mathcal{B}}_{Y,1} \ldots, e^{\mathcal{B}}_{Y,N} \right\}$ over all κ realizations.

9. For each realization, substitute the values of $\mathbf{e}^{\mathcal{B}} = \left\{ e_{Y,1}^{\mathcal{B}} \ldots , e_{Y,N}^{\mathcal{B}} \right\}$ and the estimate of the variances $\sigma_{\eta,i}^2, \ i = 1, \ldots, N$ into Eq (13) to obtain a realization of s.

12. Denote by Q_α and $Q_{1-\alpha}$ the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of the set of κ realizations of s. These are estimates of $L_{s,\alpha}^{\mathcal{B}} \ U_{s,\alpha}^{\mathcal{B}}$ respectively.

These scaled limits may serve as pre-survey quality measures for block kriging in the same way as the equivalents for the point kriging prediction intervals. In addition, we considered a further derived quality measure. Consider a case in which further investigation is required should the value of Z across a block exceed some threshold Z_t (e.g. a regulatory limit for a potential contaminant). We might choose to initiate such an investigation if the probability that the block median exceeds the threshold exceeds some value τ. In the procedure described above we could do this by finding the quantile $Q_{1-\tau}$ of the realizations of s and then investigating further if

$$Q_{1-\tau}\widehat{\widetilde{Z}} > Z_t,$$

where $\widehat{\widetilde{Z}}$ is obtained from Eq. (12). A pre-survey quality measure can be obtained by asking by what proportion, $P_{\tau,e}$ the expected block median will exceed the threshold Z_t for a block where $Q_{1-\tau}\widehat{\widetilde{Z}} = Z_t$. Simple algebra shows that this is given by:

$$P_{1-\tau,e} \ = \ \frac{1 - Q_{1-\tau}}{Q_{1-\tau}}. \tag{14}$$

The greater the uncertainty in the prediction the larger P_e will be. Depending on the application one might select a largest acceptable value, such as 0.1 or 0.2.

3. A case study with soil data

3.1. The soil data

The soil data are taken from the British Geological Survey's Geochemical Baseline Survey of the Environment (G-BASE) (Johnson et al., 2005). Specifically we used data from the Humber-Trent region, which is an area of approximately 15,800 km^2 in North

East England. The survey of this region is described in more detail by Rawlins et al (2003). In summary, the G-BASE data were collected by sampling alternate 2-km squares of the UK Ordnance Survey grid. A sample site was selected at random within each of these squares, and five soil cores were collected from the centre and corners of a 20-m square and bulked. Each core was 15-cm long and excluded surface litter. The bulked material from each site was subsequently air-dried, disaggregated, sieved to pass 2 mm, then coned and quartered. From each a 50-g sub-sample was ground in an agate planetary ball mill until 95% of the material was finer than 53 μm. The total concentrations of 26 major and trace elements were determined in each sample by wavelength dispersive XRFS (X-Ray Fluorescence Spectrometry). In this study we analysed observations from 5892 sites.

3.2. Analysis

We examined the distributions and summary statistics of the data on all 26 elements. We discarded four elements for which a substantial proportion of observations were below the detection limit. Of the remaining elements 11 had skewed distributions and for nine of these transformation to logarithms reduced the skewness substantially. We selected three of these nine elements for detailed study, Cu, Zn and As, because of their importance as indicators of soil quality. Histograms of the transformed data are presented in Figure 1. Summary statistics for the raw data and data after log-transformation are presented in Table 1. These include the median absolute deviation (MAD) of the data from their median, multiplied by a consistency correction (1.483) to provide a robust estimate of the standard deviation of the data, resistant to outlying values.

The octile skew (Brys et al., 2003) is a measure of the degree of symmetry of the 1st and 7th octiles of the data about the median. It is a robust measure of skewness which indicates the underlying symmetry of the distribution, whereas the conventional coefficient of skewness, which is computed from third and second moments of the data,

is sensitive to a few extreme values. Webster and Oliver (2007) recommend that data are considered for transformation when the conventional coefficient of skewness lies outside the interval (-1,1), and Lark et al (2006) found that a corresponding interval for the octile skew is (-0.2,0.2). Table 1 shows that all these variables have large positive coefficients of skewness on the original scales of measurement. Transformation to logs brings the octile skews close to zero in all cases. However, the conventional coefficient of skewness, while much reduced by transformation, still exceeds 1 for all three elements. The contrast between this and the results for the octile skewness suggest that there are some outlying values in the data, to which the octile skewness is resistant while the conventional skewness is not.

Exploratory geostatistical analysis showed no evidence of marked anisotropy in these variables, particularly at shorter lags where the variogram is most influential in kriging. Variograms were estimated for all three variables using the conventional method of moments estimator due to Matheron (1962) as well as three robust estimators reviewed by Lark (2000). These were considered because of the evidence for outlying data provided by the exploratory analysis described above. The robust estimators are those proposed by Cressie and Hawkins (1980), Dowd (1984) and Genton (1998).

Double spherical variogram models were fitted to the estimated variograms using the FVARIOGRAM procedure in GenStat (Payne, 2010), and weighted least squares. The double spherical model was used because of prior evidence from analysis of G-BASE data in this region that there are such nested structures attributable to effects of parent material (Rawlins et al., 2003). The Akaike Information Criterion, computed after Webster and Oliver (2007), also indicated that the double spherical model was preferable to alternatives. The variograms were then tested by cross-validation, using the XVOK3D program in GSLIB (Deutsch and Journel, 1997). This program predicts each observation in the data set in turn by ordinary kriging from the remaining data, and reports the kriging prediction, $\tilde{Y}(\mathbf{x})$ and the kriging variance $\sigma_K^2(\mathbf{x})$. We then

computed the standardized square cross validation prediction error at each location, for kriging with each variogram model. This is defined as

$$\theta(\mathbf{x}) = \frac{\left(\tilde{Y}(\mathbf{x}) - Y(\mathbf{x})\right)^2}{\sigma_K^2(\mathbf{x})}, \tag{15}$$

where $Y(\mathbf{x})$ is the observed value. Normal Q-Q plots of the cross-validation errors, in which the quantiles of the errors are plotted against the corresponding quantiles of a normal random variable, were examined, and suggested that the prediction errors appeared to be normally distributed, although possibly with some outliers. The expected value of a set of observations of $\theta(\mathbf{x})$ in circumstances where the variogram model, and so the kriging variance, is reliable is 1, and the median value is 0.455. The median is preferred as a diagnostic when considering the possibility that the data contain outlying values since it will be robust to these (Lark, 2000). We therefore followed Lark (2000) in selecting the variogram model for which the median standardized squared cross-validation error was closest to 0.455. Table 2 presents cross-validation statistics, and Figure 2 shows normal Q-Q plots for cross validation by the preferred model for each variable. The variograms estimated by Matheron's estimator and by the preferred robust estimator for each element, along with the fitted models, are shown in Figure 3.

We then considered point lognormal kriging of each element, with the selected variogram, at a target point at the centre of a square grid (i.e. at the maximum distance from any observed value), with grids of various sampling densities ranging from 4 samples km^{-2} (a 500-m square grid, the sampling density used by the British Geological Survey for many soil surveys in urban regions) to 0.01 samples km^{-2} (a 10-km grid). We computed the kriging variances at the target point for each grid and element. We then computed the corresponding upper and lower standardized limits of the prediction interval, with $\alpha = 0.1$, using Eq. (5).

We then considered lognormal block kriging of each element to a square block with sides 250 m long. The block was centred at the centre of a square grid (i.e. the block centre was at the maximum distance from any observed value). We considered the same

17

sampling densities used for point kriging, and used the procedure described in Section 2.3.3 above to calculate upper and lower standardized limits for the prediction interval of the block median with $\alpha = 0.1$. The IMSL routine DCHFAC was used to compute the Cholesky factorizations of covariance matrices, and the routine RNMVN was used to generate the realizations of the random variates (Visual Numerics, 2006). We used a discrete approximation to the block of 49 points on a 7×7 uniform grid with points on the vertices of the block. No appreciable change resulted from using more points in the discretization. We used $\nu = 1000$ randomly selected points within the block to estimate the spatial mean for each realization. We generated 50,000 realizations of the variable s in order to generate prediction intervals of the scaled block median over the model distribution.

We also computed quantile $Q_{0.25}$ of the realizations of s for each element and sampling grid, and from this computed $P_{0.25,e}$, using Eq. (14), which is the expected proportion by which the block median exceeds some threshold value for a block selected for further investigation because the computed probability that the block median does indeed exceed the threshold is 0.75.

3.3. Results

Table 2 shows the cross-validation results for the variograms obtained by various estimators. For all elements the estimator due to Matheron, which is not robust to extreme values, gave median values of θ notably smaller than the expected value of 0.455, suggesting a tendency to overestimate the prediction error variance at most locations due to the influence of outliers on the variogram. By contrast the mean value of θ for the variograms obtained with Matheron's estimator are generally close to 1, as Lark (2000) showed by simulation, this can be attributed to the combined effect of outliers on the variogram and on the cross-validation errors. We therefore selected a robust estimator for the variogram of each element, for which the median value of θ was close to 0.455. This was the estimator proposed by Cressie and Hawkins (1980)

for As and Zn, and the one proposed by Genton (1998) for Cu. The variogram models can be seen in Figure 3.

The key results are in Figure 4. Graphs are presented, for each element, of the scaled prediction interval bounds for the point-kriged estimate of the element on the original scale of measurement, or for the block median. On these graphs the sample density of the National Soil Inventory of England and Wales (NSI, a 5-km square grid) and of GBASE is shown by a vertical line. Also shown are the graphs of $P_{0.25,e}$ against sample effort, i.e. proportion by which expected median exceeds a threshold in a block where the estimated probability that the threshold is exceeded is 75%.

Figure 4 shows the following. First, we can see how the scaled prediction intervals for the point value of the variables, or the block median, become narrower as the sampling density is increased. In all cases this response to sample effort becomes rather small with sampling densities larger than about 0.2 samples per km^2. We can therefore see that the GBASE sampling scheme does allow rather more precise predictions than the NSI, but that the effect of reducing the GBASE sampling effort by, say, one third, would be rather small as judged by these results. This is of some practical relevance since the GBASE survey of the United Kingdom is not yet complete.

Second, we can see that, for point kriging, the notional quality measure of a lower bound which is at least 75% of the median cannot be achieved with the sampling densities that are considered here, and a data user would have to recognize that it was not a realistic standard for point predictions. With a sample density of 4 points per km^2 we can achieve a lower bound on the prediction interval that is 65% of the conditional median in the case of As and Zn. For kriging the median across a 250-m square block the quality measure is achievable at more manageable sample densities (just below the GBASE density for As), and at rather larger density (about 1 sample per km^2) for Cu and Zn.

Third, we can see that the quality requirement that the upper bound of the

prediction interval is no more than 1.75 times the conditional median (or block median) is much less stringent than the 0.75 standard for the lower bound. It can be achieved for As and Zn by point kriging from samples of density 0.2 points per km^2 (2.25-km grid), and from 1 point per km^2 for Cu. For block kriging of the median across a 250-m block it is achieved from a relatively coarse square grid with a spacing of about 9.5 km for Zn and somewhat larger than the maximum 10-km grid considered here for Cu and As.

Fourth, the graphs of $P_{0.25,e}$ against sample effort show that the requirement that the expected median value of a block exceeds a regulatory threshold by a factor of no more than 0.15 for a block for which the estimated probability that the median value exceeds a regulatory threshold is 0.75 can be met for a 250-m square block by sampling at about 0.25 samples per km^2 for As, at about 0.5 samples per km^2 for Cu and at about 0.9 samples per km^2 for Zn. Consider a practical example. Land managers in England and Wales are required to investigate further before applying manures to land with soil copper concentrations that exceed 80 mg kg^{-1} according to Defra (2009) (a slightly larger threshold is allowed if the soil pH exceeds 5.5). A block for which the probability that the median exceeded this threshold was 0.75 would have an expected median concentration of 92 mg kg^{-1} when kriging is done from a grid of density 0.64 samples per km^2, but would be 103 mg kg^{-1} if we predicted from the NSI grid.

4. Discussion

We have shown how pre-survey quality measures for lognormal kriging surveys of soil can be based on variograms of soil properties, transformed to the lognormal scale. We have shown how these measures might be used to compare the expected outcomes of surveys conducted with different levels of effort.

The approach requires that we consider prediction intervals for unobserved quantities, expressed as proportions of the median value of the variable at a point or across a block. We believe that this is a feasible approach, given the use of median values to

20

characterize variables with skewed distributions in standard exploratory data analysis procedures, and the difficulty of interpreting variances of skewed variables. Our approach may be criticized on two counts. First, are such scaled prediction intervals the quality measures that users require? We accept that users may commonly be interested in other quality measures, post survey, which give absolute measures of uncertainty for particular predictions. However, as we have seen, these cannot be computed pre-survey for the lognormal case. It is therefore necessary to accept, for the case of geostatistical survey of lognormally distributed variables, that the quality measures that can be used, pre-survey, for tasks such as selecting a sampling intensity, are more restricted than the quality measures that can be reported post-survey for particular predictions. Second, one might ask whether the block median which we use to develop a quality measure for block kriging is a useful target quantity. We suggest that it is in some circumstances, but not in others. An example of the latter is the original problem for which block kriging was developed, the estimation of the ore grade of a panel in a mine which the miner either considers extracting, or has to extract to access deeper material. In this case the block mean is certainly the quantity of interest, since it corresponds to the overall yield of the block and so the economic return to the effort taken to extract and process it. However, in other cases the block median may be useful. Consider the precision agriculture context, for example, in which we want to determine a fertilizer application rate for a region from a prediction of the available nutrient concentration in the soil. If we base the rate on the regional median then half the region will be somewhat overfertilized and half somewhat underfertilized. If, on the other hand, we fertilized according to the block mean then rather more than half of the region would be underfertilized. Similarly, when assessing the impact of land remediation, the block median may be a useful quantity to estimate since it would allow for more robust comparisons before and after the intervention.

In this paper we have assumed that the prediction intervals that we work with are

centred on the median of the unobserved variable in the sense that the probability that the unknown variable falls between the lower bound of the interval and the median is $\frac{1-\alpha}{2}$. It is, of course, possible to define other prediction intervals that satisfy Eq. (3), and in the case of asymmetrically distributed random variables some of these are shorter than the median-centred interval, and a shortest interval can be found (Dahiya and Guttman, 1982). De Oliveira and Rui (2009) discuss shortest prediction intervals in the case of lognormal kriging, and it would be interesting to investigate how our work on pre-survey quality measures based on prediction intervals could be extended to use bounds on the shortest interval

5. Conclusions

To conclude, the standard pre-survey quality measures used to plan geostatistical surveys are based on the kriging variance, but these cannot be applied in the case of variables which must be predicted by lognormal kriging because in this case the prediction error variance for a variable depends, inter alia, on its conditional mean. Rather, we have shown how prediction intervals for the point lognormal kriging prediction, expressed as proportions of the unknown conditional median, and similar scaled prediction intervals for the block median, can be used as pre-survey quality measures for geostatistical surveys of lognormally-distributed random variables.

Acknowledgements

References

Allègre, C.J., Lewin, E. 1995. Scaling laws and geochemical distributions. Earth and Planetary Science Letters. 132, 1–13

Bone, J., Head, M., Barraclough, D., Archer, M., Scheib, C., Flight, D., Voulvoulis, N. 2010. Soil quality assessment under emerging regulatory requirements. Environment International, 36, 609–622

Brus, D.J., Heuvelink, G.B.M. 2007. Optimization of sample patterns for universal kriging of environmental variables. Geoderma, 138, 86–95.

Brus, D.J., Noij, I.G.A.M. 2008. Designing sampling schemes for effect monitoring of nutrient leaching from agricultural soils. European Journal of Soil Science, 59, 292–303.

Brys, G., Hubert, M., Struyf, A. 2003. A comparison of some new measures of skewness. In: Developments in Robust Statistics (eds R. Dutter, P. Filzmoser, U. Gather and P.J. Rousseeuw), pp. 98–113. Physica-Verlag, Heidelberg.

Burgess, T.M. & Webster, R. 1980. Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. Journal of Soil Science 31, 315–331.

Chilès, J.-P., Delfiner, P. 1999. Geostatistics, Modeling Spatial Uncertainty. John Wiley & Sons, New York.

Cressie, N.A.C. 1993. Statistics for Spatial Data. Revised edition. John Wiley & Sons, New York.

Cressie, N. 2006. Block kriging for lognormal spatial processes. Mathematical Geology 38, 413–443.

Cressie, N., Hawkins, D. 1980. Robust estimation of the variogram. Journal of the International Association of Mathematical Geology 12, 115–125.

Dahiya, R.C., Guttman, I. 1982. Shortest confidence and prediction intervals for the log-normal. Canadian Journal of Statistics 10, 237–317.

de Gruijter, J.J., Brus, D.J., Biekens, M.F.P., Knotters, M. 2006. Sampling for Natural Resource Monotoring. Springer, Berlin.

de Oliveira, V, Rui, C. 2009. On shortest prediction intervals in log-Gaussian random fields. Computational Statistics and Data Analysis 53, 4345–4357.

Defra. 2009. Protecting our water, soil and air. The Stationary Office, Norwich.

Deutsch, C.V, Journel, A.G. 1997. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York.

Dowd, P.A. 1984. The variogram and kriging: robust and resistant estimators. In: Geostatistics for Natural Resources Characterization (eds G. Verly, M. David, A.G. Journel and A. Marechal), Part 1. pp. 91–106. D. Reidel, Dordrecht.

European Commission, 2006. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions. Thematic strategy for soil protection COM(2006) 231. Brussels, Belgium.

Di H.J., Trangmar, B.B., Kemp, R.A. 1989. Use of geostatistics in designing sampling strategies for soil survey. Soil Science Society of America Journal 53, 1163–1167.

Genton, M.G. 1998. Highly robust variogram estimation. Mathematical Geology 30, 213–221.

Johnson, C.C., Breward, N., Ander, E. L., Ault, L. 2005. G-BASE: Baseline geochemical mapping of Great Britain and Northern Ireland. Geochemistry: Exploration, Environment, Analysis 5, 1–13.

Lark, R.M. 2000. A comparison of some robust estimators of the variogram for use in soil survey. European Journal of Soil Science 51, 137–157.

Lark, R.M. 2009. Estimating the regional mean status and change of soil properties: two distinct objectives for soil survey European Journal of Soil Science 60, 748–756.

Lark, R.M, Bellamy, P.H. Rawlins, B.G. 2006. Spatio-temporal variability of some metal concentrations in the soil of eastern England, and implications for soil monitoring. Geoderma 133, 363–379.

Marchant, B.P., Lark, R.M. 2006. Adaptive sampling for reconnaissance surveys for geostatistical mapping of the soil European Journal of Soil Science 57, 831–845.

Marchant, B.P., Lark, R.M. 2007. Optimized sample schemes for geostatistical surveys. Mathematical Geology 39, 113–134.

Matheron, G. 1962. Traité de Géostatistique Appliqué, Tome 1. Mémoires du Bureau de Recherches Géologiques et Minières, Paris.

Matheron, G. 1974. Effet proportionnel et lognormalité, ou: le retour du serpent de mer. Note Géostatistique No. 124. Ecole des Mines de Paris.

McBratney, A.B., Webster, R., Burgess, T.M. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalised variables. I. Theory and Method. Computers and Geosciences 7, 331–334.

Paul, R., Cressie, N. 2011. Lognormal block kriging for contaminated soil. European Journal of Soil Science 62, 337–345.

Pawlowsky-Glahn, V., Olea, R.A. 2004. Geostatistical Analysis of Compositional Data. Oxford University Press, New York.

Payne, R.W. (editor). GenStat Release 13 Reference Manual Part 3: Procedure Library PL21. VSN International, Hemel Hempstead, Hertfordshire.

Rawlins, B.G., Webster, R., Lister, T.R. 2003. The influence of parent material on topsoil geochemistry in Eastern England. Earth Surface Processes and Landforms 28, 1389–1409.

Stein, M.L. 1999. Interpolation of Spatial Data: some Theory for Kriging. Springer, New York.

Tolosana-Delgado, R., Pawlowsky-Glahn, V. 2007. Kriging regionalized positive variables revisited: sample space and scale considerations. Mathematical Geology 39, 529–558.

van Groenigen, J.W., Siderius, W., Stein, A. 1999. Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma 87, 239–259.

Visual Numerics, 2006. IMSL Fortran Numerical Library Version 6.0. Visual Numerics, Houston, Texas.

Webster, R., Oliver, M.A. 2007. Geostatistics for Environmental Scientists. 2nd Edition John Wiley & Sons, Chichester.

White, R.E., Haigh, R.A., MacDuff, J.H. 1987. Frequency distributions and spatially dependent variability of ammonium and nitrate concentrations in soil under grazed and ungrazed grassland. Fertilizer Research. 11, 193–209.

Zhu, Z., Stein, M. 2006. Spatial sampling design for prediction with estimated parameters. Journal of Agricultural, Biological and Environmental Statistics 11, 24-44.

Table 1. Summary statistics of raw and transformed data.

	Cu	Zn	As	lnCu	lnZn	lnAs
		mg kg^{-1}			ln (mg kg^{-1})	
Mean	22.48	92.6	16.3	2.88	4.31	2.63
Median	18	73	14	2.89	4.29	2.64
SD	22.57	121.24	14.2	0.66	0.6	0.53
MAD*	8.9	34.1	5.93	0.55	0.47	0.5
Skewness	22.48	92.6	16.3	2.88	4.31	2.63
Octile skew	0.31	0.28	0.2	-0.02	0.01	-0.06
Q1	13	53	10	2.56	3.97	2.3
Q3	26	100.25	19	3.26	4.61	2.94

*Median Absolute Deviation from Median.

Table 2. Statistics on standardized squared cross validation error for each element with variograms obtained by each estimator.

Element	Variogram estimator							
	Matheron		Cressie-Hawkins		Dowd		Genton	
	Mean θ	Median θ	Mean θ	Median θ	Mean θ	Median θ	Mean θ	Median θ
As	1.057	0.265	1.734	0.443	2.018	0.519	1.793	0.471
Cu	1.039	0.289	1.504	0.421	1.76	0.489	1.715	0.473
Zn	1.064	0.27	1.699	0.432	2.079	0.521	2.106	0.534

Figure captions.

1. Histograms of log-transformed data on concentrations of (top) As, (middle) Cu and (bottom) Zn in the soils of the Humber-Trent region.

2. Q-Q plots of cross-validation prediction errors for (top) As, (middle) Cu and (bottom) Zn. The bisector is also drawn on each graph. Prediction is by ordinary kriging using the variogram estimated by the method of Cressie and Hawkins (1980) for As and Zn, and the method of Genton (1998) for Cu.

3. Variograms of transformed data on (top) As, (middle) Cu and (bottom) Zn. In each graph the solid symbol shows the estimates of the variogram obtained by the standard estimator due to Matheron, and the open symbol shows the estimates obtained by the robust estimator shown by cross-validation to provide the best variogram model. This is the Cressie and Hawkins estimator for As and Zn and Genton's estimator for Cu. A double spherical model is fitted to each empirical variogram.

4. (Left) Plot of scaled prediction interval bounds (for (solid line, and solid symbol at 4 samples km^{-2}) point kriging of element concentration on the original scale or (broken line and open symbol at 4 samples km^{-2}) block kriging estimate of the block median at the centre of square grids of different sample densities. (Right) Plot of $P_{0.25,e}$ against sample density. $P_{0.25,e}$ is the expected proportion by which the block median exceeds some threshold value for a block selected for further investigation because the computed probability that the block median does indeed exceed the threshold is 0.75. The horizontal broken line shows a target value for $P_{0.25,e}$ of 0.15.

As

Cu

Zn