Geoderma, Volume 133, Issues 3-4, August 2006, Pages 363-379

Spatio-temporal variability of some metal concentrations in the soil of eastern England, and implications for soil monitoring

R.M. Lark[a,*], P.H. Bellamy[b] , B.G. Rawlins[c]

[a]*Silsoe Research Institute, Wrest Park, Silsoe, Bedford MK45 4HS, U.K.*

[b]*National Soil Resources Institute, Cranfield University, Silsoe, Bedford, MK45 4DT, U.K.*

[c]*British Geological Survey, Keyworth, Nottingham NG12 5GG, U.K.*

Abstract.

Previous workers have proposed the use of multivariate geostatistics for the problem of estimating temporal change in soil properties for soil monitoring, but this has yet to be evaluated. We present a case study of this approach from the Humber-Trent region in North East England. We extracted data from two sources on cobalt, nickel and vanadium concentrations in the topsoil on two dates. Autovariograms were estimated for each metal on each date, and pseudo cross-variograms for each metal on the two dates. It was shown that robust estimators of the auto and pseudo cross-variograms were needed for the analysis of these data. A linear model of coregionalization was then fitted to describe the spatio-temporal variability of each metal.

While the concentration of each metal in the soil showed pronounced spatial dependence, that we know is driven by parent material, the change over time was only spatially structured for cobalt and vanadium. This shows that information on spatial variability from a single date may be a poor guide to the design of a monitoring scheme. We showed how the cokriging variance of the change in concentration of cobalt and vanadium depends on sampling effort and strategy. The change in these particular variables between two dates is best estimated by sampling with equal intensity at the same sites on both dates; and when resampling an existing baseline survey it is best to sample them at rather than

between the original sites. The best strategy in any case depends on how the variable is coregionalized over time.

Keywords: geostatistics; pseudo cross-variogram; cokriging; robust estimation; soil monitoring; heavy metals; pedometrics.

---

*Corresponding author: Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ.
E-mail address: murray.lark@bbsrc.ac.uk.

1. Introduction

1.1 The soil monitoring problem

There is considerable interest in how best to monitor the quality of the soil, so as to ensure that it is managed sustainably (e.g. Mol et al., 1998; DETR, 2001; Huber et al., 2001). There are many important questions that must be addressed if the soil is to be monitored adequately. One of these is how soil properties should be sampled in order to detect changes over time with adequate precision. This has been addressed in the context of how to estimate the mean change of a variable within an individual monitoring site (e.g. Miller et al., 2001). In this paper we are concerned with how to map changes in soil quality as spatial variables. This will be necessary in order to identify where particular problems are emerging, and where effort for improved management should be targeted.

This is a challenging problem. Various studies have shown that the spatial components of soil variation can be very large in multitemporal data sets (e.g. Bringmark and Bringmark, 1998; Webb et al., 2000). Papritz and Webster (1995a) point out that, without adequate sampling design and analysis, it may not be possible to detect important changes in soil properties because of spatial variability. They considered design-based and model-based sampling designs for soil monitoring, the latter based on a proposal of Papritz and Flühler (1994) for geostatistical mapping of change in soil variables. They then

used simulated data to illustrate how these might be implemented (Papritz and Webster, 1995b).

1.2 The coregionalization model of spatio-temporal variation

The model-based method for estimating change proposed by Papritz and Flühler (1994) treats a soil variable on m dates as a realization of m coregionalized random variables and so exploits any temporal persistence of the spatial variation that can be expressed as cross-correlation between values of the variables on different dates. When this has been done the change in the variable between two particular sampling dates can be estimated for unsampled sites, or for blocks, by cokriging, which also supplies an estimate of the estimation variance.

The advantages of this approach are twofold. First, if there is cross-correlation over time then, once an initial baseline survey has been conducted, subsequent resampling for monitoring could be done less intensively while maintaining adequate precision for estimates of the change of the soil variable at unsampled sites. Second, cokriged estimates of change are coherent, that is to say the estimate of change and both the estimates of the variable on two dates are all best linear unbiased predictors (see Webster and Oliver, 2001).

This multivariate geostatistical model of multitemporal soil data could be used to address various questions. Van Duijvenbooden (1998) points out that various European countries are developing monitoring schemes based on grid samples of the soil. For example, England and Wales has a 5-km grid baseline survey, which is likely to provide a basis for future monitoring (McGrath and Loveland, 1992). We may ask whether, having established a baseline survey, at what intensity we must resample on future dates in order to estimate changes in the soil with adequate precision. We may also ask whether it is best to attempt to resample the same sites, or to sample at other locations. Such questions could be addressed with the coregionalization model.

1.3 Alternative geostatistical approaches and their limitations

Geostatistical studies of real data, motivated by an interest in soil monitoring, have mostly focussed on analysis of data for a single time (e.g. Arrouays et al., 2000; Scholz et al. 1999). This is understandable since most soil monitoring activity is at an early stage of development. However, the limitation of these studies is that we cannot be confident that the spatial variability of a soil property on one date will be a good guide to how we should sample in order to monitor change in the soil. For example, the spatial variability of heavy metals in soil may be largely determined by geological variation (e.g. Atteia et al., 1994), and this will persist over time. The task of resampling is to detect change against the baseline sample, not to re-estimate a pattern of variation dominated by the geology. The spatial variation of change in the soil properties should determine the resampling strategy, and this may be quite different from the baseline variation, as is illustrated by the results of Sun et al. (2003).

One interesting study on temporal change of soil properties using univariate geostatistics is that of Zhang and McGrath (2004) who analysed data on the organic carbon content of soils in part of the Republic of Ireland. These had been collected on two dates. These authors used ordinary kriging to estimate organic carbon content for each date separately on a common grid, then estimated the change simply by computing the differences. The estimation variance of change was computed by adding the two kriging variances. This method discards any information that the data for one date might contribute to estimates of the property on date two, and the estimate of change cannot be regarded as the best linear unbiased prediction of change because the two ordinary kriging estimates are not coherent (Webster and Oliver, 2001). Further, the estimation variance is likely to be conservative.

Despite the disadvantages of the univariate approach, it makes no restrictive assumptions about the relationship between the variation of a soil variable on several dates (e.g. that it is consistent with a linear model of coregionalization) because it does not require modelling of the cross-covariance. Further, if the sampling sites on different dates

do not coincide then it is less straightforward to model the coregionalization than in most geostatistical problems since the model must be based on pseudo cross-variograms or the generalized cross-covariances of Künsch et al. (1997). While the coregionalization model is attractive in principle, there are possible practical problems, and we require evidence of its usefulness when applied to real data on the soil.

1.4 What is needed now

We are not aware of any studies where the approach of Papritz and Flühler (1994) and Papritz and Webster (1995a) has been applied to multitemporal data on the soil — other than a rather contrived case study by one of us (Lark, 2002a) where a subsampled data set on soil water content was used to investigate a methodological problem. The objective of the present paper is to evaluate the use of the multitemporal coregionalization model for spatio-temporal variability of soil properties which are pertinent to the monitoring of soil quality. Our case study is based on concentrations of three metals in the top soil of the Humber-Trent region of eastern England. We describe the spatio-temporal variation of these metals on two dates by a linear model of coregionalization. We then use this model to test the hypothesis that the spatial variation of change in metal content of the soil can be quite different from the spatial variation of metal content on any one date. We also demonstrate how change in the concentration of the metal can be estimated by cokriging and draw conclusions on how the soil of cognate landscapes should be sampled in order to monitor change in these variables.

2. Materials and methods

2.1 Soil sampling and Analysis

We used two sources of data in this study. The first is the National Soil Inventory of England and Wales (NSI) collected by the National Soil Resources Institute (then the Soil Survey of England and Wales), and sampled on two dates the latter less intensively. The second is data from the Geochemical Baseline Survey of the Environment (G-BASE) survey of the British Geological Survey. Since the G-BASE survey of the Humber-Trent

region, an area of approximately 15,800 square kilometres in North East England, took place at a similar time to the second NSI sampling, we have limited our study to this region that is shown in Figure 1.

The NSI baseline survey in the Humber-Trent region took place between 1979 and 1984. Sampling took place at sites on a 5-km orthogonal grid aligned with the Ordnance Survey National Grid. Details of the protocol are published elsewhere (McGrath and Loveland, 1992). Soil sampling was restricted to the uppermost 15 cm of mineral soil (or less if rock intervened), or of peat, as appropriate, i.e. litter layers were not sampled. The actual sampling depth was recorded. Twenty-five cores were taken at the nodes of a 5-m grid within a 20-m square centred on the basic 5-km grid-point. The cores were bulked and mixed well in the field, double-bagged in food-grade polythene bags, and a waterproof and rot-proof label placed between the bags. The target sample mass was 450 g of air-dried soil. In the laboratory the soil was air-dried, half the material was ground to 2-mm, then a 25-g sub-sample was taken from this by coning and quartering, and ground to $<$150 μm. This subsample was extracted with aqua regia and then analyzed for a range of metals by ICP-OES or atomic absorption spectrometry (for a few metals including vanadium).

Essentially the same protocol was followed when the NSI grid was resampled in the Humber-Trent region in 1995. The sample sites were selected at random from those used in the baseline survey so that, nationally, just under 30% of sites were resampled. There had been some development in analytical methods, but reanalysis of stored soil from the baseline sample suggested that the results were comparable for the elements that we report in this paper.

The G-BASE data were collected under a non-aligned sampling scheme. The basic strata were 2-km squares of the Ordnance Survey grid, and every second square was sampled. A sample site was selected at random within the square. At each site five soil cores were collected from the centre and corners of a 20-m square. As in the NSI survey,

the cores were 15-cm long and excluded surface litter. The five cores at each site were bulked. All samples of soil were air-dried, disaggregated, sieved to pass 2 mm, 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 24 major and trace elements were determined in each sample by wavelength dispersive XRFS (X-Ray Fluorescence Spectrometry). The G-BASE sampling in the Humber-Trent region took place in the summers of 1994, 1995 and 1996, and 6411 sites were sampled.

2.2 Combining the data

Of the NSI baseline sample points 623 lie within 10-km of a G-BASE site from the Humber-Trent survey. Of these 229 were subsequently resampled. We could use these data alone to investigate spatio-temporal variability of metals in the soil but this restricts our analysis to variation over distances of 5 km or more, the basic interval of the NSI data; and because the resampled data are relatively sparse and collected at random, they are not ideally suited to spatial analysis. To obtain information on spatio-temporal variability at finer spatial scales we decided to combine the NSI resampled data and the G-BASE data into a single set ('Date 2') for comparison with the NSI baseline survey ('Date 1').

There are obvious objections to this. Both data sets are collected on topsoil (0–15-cm) with litter excluded, and both from a 20-m square, but there are differences in sample support. The NSI sample is 25 cores bulked from across the square while the G-BASE data are 5 cores bulked from the corners and centre of the square. The analytical methods are also different. The NSI is based on aqua regia extraction and ICP-OES while the G-BASE is a solid phase analysis by XRFS that determines the total concentration of an element in the soil. Experience suggests that analysis by XRFS will usually yield larger total concentrations of certain metals in soil in comparison to an aqua regia digest followed by ICP-OES, because during the acid digestion some of the more recalcitrant soil minerals may not be dissolved. This may also account, in part, for the smaller analytical variances (from the analysis of duplicates and sub-samples) for XRFS (see Table 1) in

which some of the metals may be present in different forms.

Ideally we would not combine these two data sets. Rather we would undertake a new sampling campaign or reanalyze retained soil samples with a common technique. In practice we must make the best use of available data that represents a substantial investment of resources in order to make informed decisions about how best to monitor these variables in future.

We compared the statistics of the NSI resampled data from the Humber-Trent region and the G-BASE data. In this paper we focus in detail on the analysis of data on cobalt, and report some results for nickel and vanadium. In Table 1 we present for comparison the statistics on these variables in the two data sets, and in Figure 2 we show the empirical cumulative frequency distributions. As well as standard descriptive statistics we present some robust measures of location (the median), variability (Rousseeuw and Croux's (1992,1993) Qn that estimates the standard deviation) and skew (the octile skew, see Brys et al. 2003). These statistics are resistant to the effects of outlying data that are expected in data on metal concentrations in the soil, since some point pollution is likely.

The estimator Qn is based on a linear combination of order statistics. It was proved by Rousseeuw and Croux (1992,1993) to be both robust and efficient.

$$Qn \;=\; 2.219\{|X_i - X_j|; i < j\}_{\binom{H}{2}}, \tag{1}$$

where $X_1, \ldots, X_n$ are the n ordered data, $\{\}_k$ is the kth order statistics of the terms in brackets and H is the integer part of $(n/2)+1$. The constant is a consistency correction so that Qn estimates the standard deviation if the data are drawn from a normal distribution.

The octile skew is defined as

$$\frac{(P_{0.875} - P_{0.5}) - (P_{0.5} - P_{0.125})}{P_{0.875} - P_{0.125}} \tag{2}$$

where P_q is the value of the ordered datum such that proportion q of the data are smaller than P_q (Brys et al., 2003). The octile skew is zero if the 1st and 7th octiles are symmetric about the median. Data with a conventional coefficient of skew larger than 1.0 are usually

transformed (Webster and Oliver, 2001). We found that random variables drawn from distributions in Tukey's g family with a conventional coefficient of skew of 1.0 have an octile skew close to 0.2.

The three variables shown in Table 1 and Figure 2 are among those where we combined the NSI resampled data with the G-BASE data. In all cases the robust measures of skew and dispersion were of similar order and the shapes of the empirical cumulative distribution functions were also similar. The distributions differed more or less in their means and medians, which was attributed to differences between the analytical methods so in all cases the data sets were adjusted to a common median value (that of the set with the largest value). This was done by an additive correction, i.e. by adding the absolute difference between the medians of the two sets to each value in the set with the smaller median. Note that the median and mean values of the combined data set do not affect the LMCR that is fitted, since this depends only on the variances and covariances of the values on the two dates.

2.3 Spatial Analysis

We assume that the value of a soil property at time u and location $\mathbf{x}$, $z_u(\mathbf{x})$ is a realization of an intrinsically stationary random function $Z_u(\mathbf{x})$. Intrinsic stationarity is discussed in detail elsewhere (e.g. Webster and Oliver, 2001). It includes the assumption that we may define an auto-variogram function $\gamma_{u,u}(\mathbf{h})$ where $\mathbf{h}$ is a spatial interval, the lag. If the soil is observed on a second date, v, then a new random function, $Z_v(\mathbf{x})$, is invoked with auto-variogram $\gamma_{v,v}(\mathbf{h})$. The spatial covariation of the variable on the two dates is described by a cross-variogram function:

$$\gamma_{u,v}(\mathbf{h}) \;=\; \frac{1}{2}\mathrm{E}[\{Z_u(\mathbf{x}) - Z_u(\mathbf{x}+\mathbf{h})\}\{Z_v(\mathbf{x}) - Z_v(\mathbf{x}+\mathbf{h})\}]. \tag{3}$$

The auto-variograms and cross-variograms can be estimated for different lags, then a model is fitted to the estimates. The joint model of the variograms must have certain properties so that the overall coregionalization model is positive definite (all combinations of the constituent random variables have a positive variance). This is most readily ensured

by fitting a linear model of coregionalization (LMCR). We do not discuss this here, details are given by Goovaerts (1997) and by Lark and Papritz (2003), but it does impose the requirement that the values of the variable on different dates can be regarded as linear combinations of a common set of random variables.

When a LMCR has been obtained for two variables then their values, or linear combinations such as the difference $z_v(\mathbf{x}_0) - z_u(\mathbf{x}_0)$, at unsampled sites, $\mathbf{x}_0$, or over blocks can be estimated by cokriging. Cokriging is an optimal estimator in the sense that the mean square estimation error is minimized, an estimate of this error is also provided (the cokriging variance). This variance is determined by the disposition of sample points around the target location or block where the estimate is made, and the LMCR. It is therefore possible, once an LMCR is obtained, to compare different sampling strategies with respect to the cokriging variances and so to select a strategy that is both economical and fit for purpose (McBratney and Webster, 1983).

This is not straightforward in the present case. First, we cannot use all available data to estimate the cross-variogram since this can only use data from sites where all variables are measured. Second, data on soil metal concentrations over a large area such as this are likely to contain outliers from contaminated hot spots. These will inflate variogram estimates (Lark, 2000) and so will give misleading information on the precision with which can be estimated from a given sample. We now discuss these issues in more detail.

2.3.1 The Pseudo cross-variogram. In this case study only the resampled NSI sites in the Date 2 data set are at locations common to Date 1, that is 229 sites. This is a small proportion of all the data available and no information on spatial variation over distances less than 5 km is available from these sites, but it is only these data that may be used directly to estimate the cross-variogram. For this reason we used an alternative variogram, the pseudo cross-variogram defined by Myers (1991) as:

$$\gamma^{\mathrm{P}}_{v,u}(\mathbf{h}) \;=\; \frac{1}{2}\mathrm{Var}[Z_v(\mathbf{x}+\mathbf{h}) - Z_u(\mathbf{x})], \tag{4}$$

where Var[ ] denotes the variance of the term in brackets. An alternative approach is to follow Künsch et al. (1997) and to fit an LMCR to generalized cross-covariances. Here we used the pseudo cross-variogram since suitable robust estimators exist (see the discussion below).

Papritz et al. (1993) describe the pseudo cross-variogram in detail. Here we note only that, while the pseudo cross-variogram can be defined for any pair of weakly stationary random variables it only exists in certain conditions when the variables are only intrinsically stationary.

A LMCR may be fitted to estimates of the auto-variogram and pseudo cross-variogram of two variables; see, for example, Lark (2002a). Certain assumptions are necessary if an estimate of the pseudo cross-variogram at lag zero is not available, (Papritz et al., 1993; Lark, 2002a) but this is not the case in the present study. When an LMCR has been fitted, then the cokriging estimates of linear combinations of the constituent variables can be obtained, as described by Papritz and Flühler(1994).

2.3.2 Robust estimators of the pseudo cross-variogram. A standard estimator of the pseudo cross-variogram is the centred estimator proposed by Papritz et al. (1993):

$$\widehat{\gamma}_{v,u}^{\mathrm{P,Pa}}(\mathbf{h}) \;=\; \frac{1}{2N_{v,u}(\mathbf{h})} \sum_{i=1}^{N_{v,u}(\mathbf{h})} [\{z_v(\mathbf{x}_\mathrm{i}+\mathbf{h}) - \bar{z}_v\} - \{z_u(\mathbf{x}_\mathrm{i}) - \bar{z}_u\}]^2, \tag{5}$$

where $\bar{z}_u$ and $\bar{z}_v$ are the arithmetic averages of the variable z on dates u and v.

This estimator is equivalent to Matheron's (1962) estimator of the auto-variogram when $v = u$, and like Matheron's estimator it is very susceptible to the effect of a few outlying values. Outliers can be values unusually large or small relative to the marginal distribution of the data ('marginal outliers'), or values that are only unusual relative to those at neighbouring locations (Laslett's & McBratney's (1990) 'spatial outliers'). Outliers, other than simple errors, arise from processes different from those that cause soil variation over most of the region. For example, deposits of dung and urine on pasture give patches of soil with larger nutrient concentrations than the background concentrations elsewhere (McBratney & Webster 1986).

Outliers may represent important phenomena that we wish to detect, such as 'hot spots' of pollutants, but they should not influence unduly the variograms used to krige over a whole site, or to design a grid for detailed sampling. Because the differences in Equation (5) are squared the effect of an outlier on the estimates is large; and, because one observation is likely to appear in several differences over different lags, one outlier can inflate the whole variogram. This is clearly undesirable.

Large errors of measurement or transcription are usually detected and removed when data are edited, but not all outliers are unambiguously wrong. If we remove large or small but legitimate values we may will bias later inferences from data. For this reason robust estimators should be used to estimate parameters from data that include outliers. Such estimators are resistant to the effects of extreme values. Lark (2000) reviewed and demonstrated some robust estimators of the auto-variogram and later showed (Lark, 2002a) that these can be generalized to robust estimators of the pseudo cross-variogram since the latter is a univariate variance.

The sample average of a data set, used in the pseudo cross-variogram estimator in Equation (5) is susceptible to outliers while the median is a robust estimator of the location of a distribution. The robust estimators of the pseudo cross-variogram first centre the data for each variable by the sample median. For simplicity of notation we define a centred difference variable:

$$Y^i_{\tilde{v},\tilde{u}}(\mathbf{h}) \equiv \{z_v(\mathbf{x}_i+\mathbf{h}) - \tilde{z}_v\} - \{z_u(\mathbf{x}_i) - \tilde{z}_u\} \tag{6}$$

where $\tilde{z}_u$ and $\tilde{z}_v$ are the sample medians.

We may generalize the robust auto-variogram estimator of Cressie & Hawkins (1980) to a pseudo cross-variogram estimator $\widehat{\gamma}^{\mathrm{P,SRD}}(\mathbf{h})$:

$$2\widehat{\gamma}^{\mathrm{P,SRD}}_{v,u}(\mathbf{h}) = \frac{\left\{\frac{1}{N_{2,1}(\mathbf{h})}\sum_{i=1}^{N_{v,u}(\mathbf{h})}|Y^i_{\tilde{v},\tilde{u}}(\mathbf{h})|^{\frac{1}{2}}\right\}^4}{0.457+\frac{0.494}{N_{v,u}(\mathbf{h})}+\frac{0.045}{N^2_{v,u}(\mathbf{h})}}. \tag{7}$$

The denominator is a consistency correction, on the assumption that the data uncontaminated by outliers are drawn from a bivariate normal spatial process. This estimator

has useful properties. Lark (2002a) showed that it was both resistant to outliers and efficient. However, unlike some estimators it is not B-robust, i.e. the effect of a very large contaminant is not bounded. One B-robust estimator of variance is the Median Absolute Difference. Lark (2002a) proposed the estimator:

$$2\widehat{\gamma}_{v,u}^{\mathrm{P,MAD}}(\mathbf{h}) \;=\; 2.198\{\mathrm{Median}[|Y_{\tilde{v},\tilde{u}}^{i}(\mathbf{h})|]_{i=1}^{N_{2,1}(\mathbf{h})}\}^{2}. \tag{8}$$

When $v = u$ this is equivalent to Dowd's (1984) robust variogram estimator. There are bounds on the effect of any outlier on this estimator (see Lark 2000 for discussion) and simulation showed it to be very robust although less efficient than alternatives (Lark, 2002a).

We have already used Rousseeuw & Croux's (1992, 1993) Qn as a robust estimator of the standard deviation. This may be applied to the difference variable in order to return a robust variogram estimate:

$$2\widehat{\gamma}_{v,u}^{\mathrm{P},Qn}(\mathbf{h}) \;=\; \left[2.219\{|Y_{\tilde{v},\tilde{u}}^{i}(\mathbf{h}) - Y_{\tilde{v},\tilde{u}}^{j}(\mathbf{h})|; i<j\}_{\binom{H_{\mathbf{h}}}{2}}\right]^{2}. \tag{9}$$

where $H_{\mathbf{h}}$ is the integer part of $(1 + N_{v,u}(\mathbf{h})/2)$ and $N_{v,u}(\mathbf{h})$ is the number of centred pair difference observations over lag $\mathbf{h}$ in our data. When $v = u$ this is equivalent to Genton's (1998) robust estimator of the auto-variogram.

In this study we used the standard auto-variogram estimator, $\widehat{\gamma}_{v,u}^{\mathrm{P,Pa}}(\mathbf{h})$, $\widehat{\gamma}_{v,u}^{\mathrm{P,SRD}}(\mathbf{h})$, $\widehat{\gamma}_{v,u}^{\mathrm{P,MAD}}(\mathbf{h})$ and $\widehat{\gamma}_{v,u}^{\mathrm{P},Qn}(\mathbf{h})$ to estimate auto- and pseudo cross-variograms for data on soil metal concentrations from the two dates.

Lark (2000) showed that differences between robust and non-robust estimators of the auto-variogram may reflect the presence of outliers in the data, but may also arise when our data do not resemble a contaminated normal process (a log-normal process may be a more appropriate model). It was found that we can distinguish these situations by cross-validation of the variogram model. A model is fitted to each set of estimates of the auto-variogram. Each observation, $z_u(\mathbf{x}_i)$, is then excluded from the data set in turn and estimated from the rest by ordinary kriging. This returns an estimate, $\widehat{Z}_u(\mathbf{x}_i)$ and a

kriging variance $\sigma^2_{\mathrm{K}}(\mathbf{x}_i)$. From these we may compute a standard square kriging error:

$$\theta(\mathbf{x}_i) \;=\; \frac{\left\{\widehat{Z}_u(\mathbf{x}_i) - z_u(\mathbf{x}_i)\right\}^2}{\sigma^2_{\mathrm{K}}(\mathbf{x}_i)}. \tag{10}$$

If the variogram model is correct then the expected value of $\theta(\mathbf{x}_i)$ is 1, since the squared error should equal the kriging variance on average. If the variogram is overestimated (due to outliers, for example), then the ratio will tend to be smaller than 1. However, the average value of $\theta(\mathbf{x}_i)$ over all observations will be a misleading diagnostic statistic because it is itself susceptible to the effects of outliers. Lark (2000) proposed the median value, $\tilde{\theta} \;=\; \text{median}\,[\theta(\mathbf{x}_i)]_{i=1}^{N}$ as a better diagnostic because of its robustness.

If the variogram model is correct then the expected value of $\tilde{\theta}$ is 0.455 when the background process is normal. Lark (2002b) showed how confidence limits for the sample value of $\tilde{\theta}$ could be obtained by bootstrapping to allow for correlation among the kriging errors. It is proposed that if the value of $\tilde{\theta}$ obtained with the variogram based on a non-robust estimator is significantly smaller than 0.455, then that robust estimator is used for which the value is closest to the expectation.

2.4 Analytical protocol

In this study we estimated auto-variograms for the concentration of a metal in each of the two data sets, Date 1 and Date 2. We cross-validated each model and computed confidence limits for the sample $\tilde{\theta}$ by bootstrapping. If $\tilde{\theta}$ for the auto-variogram obtained by the standard estimator was not significantly different from 0.455 on both dates, then this estimator and $\widehat{\gamma}^{\mathrm{P,Pa}}_{v,u}(\mathbf{h})$ were used to obtain auto-variograms and pseudo cross-variograms respectively. Otherwise the robustly estimated auto-variogram model that performed most consistently on both dates was chosen and the pseudo cross-variogram was estimated by the corresponding robust estimator:— $\widehat{\gamma}^{\mathrm{P,SRD}}_{v,u}(\mathbf{h})$, $\widehat{\gamma}^{\mathrm{P,MAD}}_{v,u}(\mathbf{h})$ or $\widehat{\gamma}^{\mathrm{P},Qn}_{v,u}(\mathbf{h})$.

A LMCR was then fitted to the set of estimates by weighted least-squares with the usual constraints on the coefficients to ensure a positive definite model. A simulated annealing algorithm was used that allows all model parameters to be optimized simultaneously while meeting these constraints (Lark and Papritz, 2003).

The LMCR was then used to map metal concentrations on both dates and the change variable with the cokriging equations of Papritz and Flühler (1994). The LMCR was then used to determine the following.

1. The auto-variogram of the change variable $Z_v(\mathbf{x}) - Z_u(\mathbf{x})$ where v denotes the second date. The change auto-variogram, $\gamma_{v-u}(\mathbf{h})$ is :

$$\gamma_{v-u}(\mathbf{h}) \;=\; \gamma_v(\mathbf{h}) + \gamma_u(\mathbf{h}) - 2\gamma_{v,u}(\mathbf{h}), \tag{11}$$

so we can estimate the variogram of change from the terms of the fitted LMCR.

2. The point cokriging variance for estimates of the change variable $Z_v(\mathbf{x}) - Z_u(\mathbf{x})$ from three notional square-grid sampling schemes applied at different intensities. In each case the cokriging variance was estimated for a location at the centre of a Date 1 grid cell cokriging from the nearest 36 observations on each date. The sampling schemes were as follows:

i. Sample effort (measured as samples km^{-2} over both dates) equally divided between two dates on a square grid with sampling at the same location on each date.

ii. Sample effort on Date 1 is constrained to the NSI baseline survey (a 5-km interval square grid) and sampling on Date 2 is at intervals $5/n$ km where n is an integer and the sample grids on both dates are in phase, thus all Date 1 sample sites are resampled on Date 2. Here our sampling strategy is to resample all sites on the baseline survey and to intercalate additional points.

iii. As (ii) but the Date 2 sampling is done on square grids of interval i=2,3 or 4-km so that Date 2 points only coincide with proportion p of the Date 1 points where $p = \min\left\{\frac{1}{i}, \frac{1}{i \bmod(5)}\right\}$. Here we only resample a proportion of the baseline survey sites and collect other samples on Date 2 from sites not previously surveyed.

2.5 Hypothetical examples

The cokriging variances obtained above will reflect the particular coregionalization struc-

ture of our data sets. To investigate the sensitivity of the choices between sampling strategies to the strength of the cross-correlation between the two dates we compared strategy (i) above with a fourth strategy for hypothetical LMCR. The strategy was

iv. Sample effort is divided equally between two dates on square grids both of interval i. The The grids on the same date are aligned but offset along the rows and columns by $i/2$. In practice the sampling on the second date would be supplemented by resampling some of the date 1 sites to support estimation of the pseudo cross-variogram at lag zero, but we have not considered the impact of these points on kriging here.

We computed the maximum point cokriging variance of the difference between the two dates under strategy (i) and (ii) with the same grid intervals, kriging from the nearest 36 observations for each date. We assumed an LMCR with all nugget terms zero and a spherical component of range 5 km. The sill variances for each date were set to 1.0 and three values for the covariance of the spherical components were considered: 0.0, 0.4 and 0.8.

3. Results

Table 1 shows the summary statistics for the data sets on three metals. Note that in most cases the standard moment-based coefficient of skew was larger than 1 while the corresponding robust Octile skew was smaller than 0.2. This suggests the presence of outlying values among the data. The auto-variograms for soil cobalt concentration by all estimators on both dates are shown in Figure 3. On Date 1 the difference between the non-robust estimator and the robust alternatives is particularly pronounced. Table 2 shows the cross-validation results for cobalt. Note that on both dates the median standard squared kriging error, $\tilde{\theta}$, obtained with the non-robust variogram estimator is significantly smaller than expected, suggesting that this has been overestimated due to outliers. On both dates $\tilde{\theta}$ for cross-validation of the variogram model fitted to estimates obtained from $\widehat{\gamma}_{v,u}^{\mathrm{P,SRD}}(\mathbf{h})$

were not significantly different from the expected value, and were in fact very close to it. This estimator was therefore selected to estimate the pseudo cross-variograms. These estimates and the fitted LMCR are shown in Figure 4, the model parameters are presented in Table 3. The fitted model is somewhat smaller than the estimates of the pseudo cross-variogram, this reflects the constraints on the LMCR to ensure positive definiteness, and may be due to some non-linearity in the process of change in the variable. Note that the covariance structures for the nugget terms of Cobalt and Vanadium are degenerate, that is to say the implied correlation is 1.0 This probably reflects some non-linearity in the coregionalization over short distances, and hence a poor fit of the LMCR. The spatially correlated components for Nickel also have a degenerate covariance.

Note that in the pseudo cross-variograms of cobalt and vanadium there is evidence for a very short range structure, and the fitted model overestimates the pseudo cross-variogram at the second shortest lag (about 140 m, the shortest lag for the pseudo cross-variogram is zero). This cannot be fitted by the LMCR because we do not have evidence in the Date 1 autovariogram for structure at these distances as the shortest lag there is 5 km. The LMCR can only accommodate a structure in the cross-covariance model that also appears in the auto-covariance model. However, the model fit is good at all longer lags, so our conclusions about (point) kriging variance are unaffected.

For economy of space we do not present all the details on the analysis for the other two metals, but in both cases the variogram estimates obtained with $\widehat{\gamma}_{v,u}^{\mathrm{P,SRD}}(\mathbf{h})$ were selected by the same process. The robust estimates and fitted model are shown in Figures 5 and 6, and the parameters of the LMCR are in Table 3. For all these metals there is strong spatial structure in the variability of concentrations in both dates and strong structure of the covariation revealed by the pseudo cross-variograms. Spatial dependence is seen at lag distances of up to 15 to 25 km. This is not surprising. Rawlins et al. (2003) have shown that the variation of trace metals in the soil of this region is strongly influenced by the spatial pattern of parent material, and this factor is unchanged over

time.

The variograms of the difference variable were computed from the fitted LMCR with Equation (11). There is some question over how reliable these are because of the sensitivity of the semi-variance of the change at shorter lags to the covariance structure of the nugget component of the LMCR, which is degenerate for cobalt and vanadium and does not reflect all the information in the pseudo cross-variogram because of the absence of short-range information for date 1. Equation (11) is therefore likely to underestimate the nugget component of the change variograms. As a check we computed the change variogram directly from the observed change at the relatively few resampled NSI sites from date 2. There are only 203 of these for cobalt and nickel and 155 for vanadium. Further, as we have seen, the variances of the date 2 observations from the NSI resampling are not always similar to those from the GBASE data for date 2. The comparison of the variogram estimates with the models obtained from Equation (11) must therefore be made with some caution. Both are presented in Figure 7. There is good agreement for nickel and vanadium, (although the sparsity of data on the latter means that we cannot look at lags less than 10km). The agreement for cobalt is poor, particularly at longer lags. Both sets of variograms show that the spatial structure of the change process, which is the variable of central interest in soil monitoring, may not be inferred directly from observation that the variables on either of the two dates have spatial structure. The change variables for cobalt and vanadium both show strong spatial structure, but that for nickel is close to a nugget process, essentially a flat variogram. This indicates that, in the case of nickel, the spatially structured variation is very similar on both dates, and the difference is a process of uncorrelated noise. For the other elements there is some spatial structure to the process of change in the metal concentration. This will have implications for our interpretation of the data (the likely processes driving change in each case) and for monitoring. It would be futile to try to map change in nickel concentration in this or cognate environments by kriging, for example, given the lack of spatial dependence.

The change in the other elements can be mapped, however, as long as we can sample at intervals of 7 km or less. This is illustrated in Figures 8 and 9. The former shows cokriged estimates of the metal concentration for both dates. In Figure 9 (top) is the kriged map of change in the variables. Note that this has been scaled to a mean change of zero (since we are not certain that the overall difference between the dates is due to anything other than differences in laboratory methods). This does not affect the kriging variance (lower map) that shows how the uncertainty of the map of change varies with the density of observations over the region.

The kriging variances for difference sampling schemes are shown for cobalt and vanadium in Figure 10. The kriging variance falls with sampling effort, so in principle we could read off these graphs the sample effort needed to map change in the variables to a specified level of precision. Note that the cokriging variance at a given level of effort is always smallest for case (i) where the sampling effort is divided equally between the dates. The response to extra effort is much smaller when this is concentrated on a single date. Similarly, at a given total sampling effort, the cokriging variance is smaller when the Date 2 sample includes all the Date 1 sites (dotted line), although the difference depends on the spatial variation of the change process and is bigger for cobalt than for vanadium, since the change process in the latter case has a very small nugget variance.

The cokriging variances for the hypothetical case studies are shown in Figure 11. Note that the cokriging variances decrease as the correlation between the spatially structured components of the variables on the two dates increase. Note also that strategy (iv) is preferred to (i) when the correlation is weak or absent, but it is best to resample at the original sites (as in our LMCR for cobalt and vanadium) when the correlation is strong. When the correlation is small the cokriging variance will approach the sum of the ordinary kriging variances for the two dates separately, and it is clear that this can be minimized over points within a sampled region if we minimize the distance to the nearest neighbouring sample point, regardless of the date on which it is sampled. By contrast,

when the correlation is strong we benefit from direct observation of the change variable..

5. Conclusions

We have estimated auto-variograms for metal concentrations in the soil on two different dates, and estimated the pseudo cross-variograms. For each metal these variograms could be fitted reasonably well by a linear model of coregionalization, as first proposed by Papritz and Flühler (1994). We could then use these models to draw the following conclusions that are pertinent to the problem of soil monitoring.

i. As hypothesized above, the spatial variability of the change in a soil variable, which will determine the efficiency of a particular sample design for monitoring, is not necessarily related to the variability of the property on a particular date. The spatial variability of metals in the soil of this region is driven primarily by parent material, which determines a strong spatial pattern distinct from the factors that drive change (Rawlins et al., 2003). We should be wary of planning monitoring from observations of a variable on a single date.

ii. For some variables the spatial variation of the change process means that it is possible to map this change by cokriging from a feasible sample density, while this is not possible for other variables (nickel in this case) A LMCR from a reconnaissance resampling of the baseline grid is necessary in order to identify which properties could be monitored in this way.

iii. The most efficient sampling design for cokriging estimates of change in the variables we have looked at here is to divide sampling effort equally between dates. In practice, however, we require a good baseline survey, and subsequent resampling will have to be designed within cost constraints.

iv. When we are resampling a baseline survey of the variables that we have examined here, with quite strong correlation of the spatially dependent components of variation on the two dates, then it is important to resample the original sites in order

to avoid inefficiency, particularly if the short-range variability is large. In practice resampling a site may be constrained by the disturbance caused on previous visits, and the accuracy with which the sample site can be relocated. In this case study a bulk sample from within a sampling area was used which smooths local variation, reduces the effect of relocation error, and would allow the sampler to avoid obviously disturbed microsites.

v. By contrast, the hypoithetical cases show that if the correlation between observations on two days is weak then resampling the date 1 sites on date 2 will be inefficient (although some resampling will be needed to estimate the pseudo cross-variogram at lag zero). This underlines our observation under (ii) above that a reconnaissance resampling of a baseline survey, so as to estimate the LMCR, will allow a rational decision to be made about how best to undertake a full second sampling to monitor change in soil against a baseline.

We now examine the broader implications of these conclusions for soil monitoring strategy. The first decision to be made when planning a monitoring scheme must be whether we require local or regional estimates of change in the soil. The latter may be satisfactory for many purposes, but local estimation will be necessary if one goal of monitoring is to identify where remediation may be needed, or possible sources of problems. When regional estimates are required then design-based sampling according to some scheme of randomization is straightforward and cost-effective (Papritz and Webster, 1995a). The considerations on which design-based or systematic sampling would be chosen for a problem are discussed by Brus and de Gruitjer (1997). An important point that they emphasize is that design-based sampling and estimation can be correctly applied to spatially dependent variables. It may also be chosen over geostatistical estimation when there is no evidence of spatial dependence, or when spatial dependence is only seen over distances that are too small to be resolved with available sampling effort.

If we decide that local estimation is required, then we face two questions.

i. Is it feasible to obtain local estimates of change, given the spatial variability and the available resources for sampling?

ii. How should the sampling be done to estimate change with adequate precision?

We have seen that the answer to this question depends quite sensitively on the spatial coregionalization of the variable with itself over time, and that the spatial variability of the variable on a single date may be a poor guide to this. In practice, when sampling against some existing baseline survey, we could conduct a reconnaissance sample to estimate an LMCR. This would include some resampling of previous sites, and sampling at intervening sites to give information on short-range variability. The resulting LMCR can then be used to investigate the cokriging variances for change in the variable under different sampling schemes, as we have done here.

If the LMCR indicates that local estimates of change in the variable cannot be obtained with available resources for resampling then we are confined to regional estimation. When the baseline survey is a systematic grid then this lack of randomness must be accounted for in the resampling and estimation of a regional mean. One approach would be to use cokriging of the regional mean, as discussed by Papritz and Webster (1995a).

Our discussion above assumes an essentially static approach to sampling, with the aim of estimating change relative to some baseline survey. Thus may be suitable for many problems in soil monitoring where changes are relatively small and slow. For more volatile properties, where the changes are large relative to the baseline values, an adaptive approach to monitoring may be favoured, as used by Wikle and Royle (1999) for atmospheric monitoring.

As a final comment, we accept that the LMCR may prove too restrictive a model for multitemporal soil variation. We suggest that the development of ideas for soil monitoring should start from the proposals of Heuvelink and Webster (2001) for the modelling of soil variation in space and time. In short, rather than looking for more complex statistical models we should aim to incorporate physical knowledge of the processes and mechanisms

of soil pollution or degradation into the monitoring scheme.

Acknowledgements

This research was funded by the UK Department for Environment and Rural Affairs (Defra) under grant SP0520. This paper is published with the permission of the Director of the British Geological Survey (NERC). We are grateful to Dr Andreas Papritz and to an anonymous referee for helpful comments on the first version of this paper.

References

Arrouays, D., Martin, S., Lepretre, A., Bourennane, H. 2000. Short-range spatial variability of metal contents in soil on a one hectare agricultural plot. Communications in Soil Science and Plant Analysis 31, 387–400.

Atteia, O., Dubois, J.-P., Webster, R. 1994. Geostatistical analysis of soil contamination in the Swiss Jura. Environmental Pollution 86, 315–327.

Bringmark, E., Bringmark, L. 1998. Improved soil monitoring by use of spatial patterns. Ambio 27, 45–52.

Brus, D.J., de Gruijter, J.J., 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with Discussion). Geoderma 80, 1–59.

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.

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

DETR 2001. The Draft Soil Strategy for England — a Consultation Paper. Department of Environment, Transport and the Regions (U.K.), London.

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.

van Duijvenbooden, W. 1998. Soil monitoring systems and their suitability for predicting delayed effects of diffuse pollutants. Agriculture, Ecosystems and Environment 67, 189–196.

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

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Heuvelink, G.B.M.,Webster, R., 2001. Modelling soil variation: past, present and future. Geoderma 100 269–301.

Huber, S, Syed, B., Freudenschuß, A., Ernstsen, V., Loveland, P. 2001. Proposal for a European soil monitoring and assessment framework. Technical Report 61, European Environment Agency, Copenhagen.

Künsch, H.R., Papritz, A. Bassi, F. 1997. Generalized cross-covariances and their estimation. Mathematical Geology 29 779–799.

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. 2002a. Robust estimation of the pseudo cross-variogram for cokriging soil properties. European Journal of Soil Science 53, 253–270.

Lark, R.M. 2002b. Modelling complex soil properties as contaminated regionalized variables. Geoderma 106, 171–188.

Lark, R.M., Papritz, A. 2003. Estimating parameters of the linear model of coregionalization by simulated annealing. Geoderma 115, 245–260.

Laslett, G.M., McBratney, A.B. 1990. Further comparison of spatial methods for predicting soil pH. Soil Science Society of America Journal 54, 1553–1558.

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

McBratney, A.B., Webster R. 1983. Optimal interpolation and isarithmic mapping of soil properties. V. Coregionalization and multiple sampling strategy. Journal of Soil Science 34, 137–162.

McBratney, A.B., Webster, R. 1986. Choosing functions for semi- variograms of soil properties and fitting them to sampling estimates. Journal of Soil Science 37, 617–639.

McGrath, S.P., Loveland, P.J. 1992. The Soil Geochemical Atlas of England and Wales. Blackie Academic and Professional, Glasgow.

Miller, J.D., Duff, E.I., Hirst, D., Anderson, H.A., Bell, J.S., Henderson, D.J. 2001. Temporal changes in soil properties at an upland Scottish site between 1956 and 1997. The Science of the Total Environment 265, 15–26.

Mol, G., Vriend, S.P., van Gaans, P.F.M. 1998. Future trends, detectable by soil monitoring networks? Journal of Geochemical Exploration 62, 61–66.

Myers, D.E. 1991. Pseudo-cross variograms, positive-definiteness and cokriging. Mathematical Geology 23, 805–816.

Papritz, A., Flühler, H. 1994. Temporal change of spatially autocorrelated soil properties: optimal estimation by cokriging. Geoderma 62, 29–43.

Papritz, A., Webster, R. 1995a. Estimating temporal change in soil monitoring: I. Statistical theory. European Journal of Soil Science 46, 1–12.

Papritz, A., Webster, R. 1995b. Estimating temporal change in soil monitoring: II. Sampling from simulated fields. European Journal of Soil Science 46, 13–27.

Papritz, A., Künsch, H.R., Webster, R. 1993. On the pseudo cross-variogram. Mathematical Geology 25, 1015–1026.

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.

Rousseeuw, P.J., Croux, C. 1992. Explicit scale estimators with high breakdown point. In: L_1 Statistical Analysis and Related Methods (ed. Y. Dodge), pp. 77–92. North Holland, Amsterdam.

Rousseeuw, P.J., Croux, C. 1993. Alternatives to the median absolute deviation. Journal of the American Statistical Association 88, 1273–1283.

Scholz, M., Oliver, M. A., Webster, R., Loveland, P. J., McGrath, S. P. 1999. Sampling to monitor soil in England and Wales. In geoENV II: Proceedings 2nd European Conference on Geostatistics for Environmental Applications, Valencia, 18–20 November 1998 J. Gomez-Hernandez, A. Soares and R. Froidevaux, eds, 465–476. Kluwer Academic, Dordrecht.

Sun, B., Zhou, S., Zhao, Q. 2003. Evaluation of spatial and temporal changes of soil quality based on geostatistical analysis in the hill region of subtropical China. Geoderma 115, 85–99.

Webb, K.T., Wang, C., Astatkie, T., Langille, D.R. 2000. Spatial and temporal trends in soil properties at a soil quality benchmark site in central Nova Scotia. Canadian Journal of Soil Science 80, 567–575.

Webster, R., Oliver, M.A. 2001. Geostatistics for Environmental Scientists. John Wiley and Sons, Chichester.

Wikle, C.K., Royle, J.A. 1999. Space-time dynamic design of environmental monitoring networks. Journal of Agricultural, Biological and Environmental Statistics 4, 489–507.

Zhang, C., McGrath, D. 2004. Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of southeastern Ireland from two different periods. Geoderma 119, 261–275.

Table 1 Summary statistics for three elements in all three data sets along with analytical variances and detection thresholds determined in the respective laboratories. Units are mg kg^{-1}.

	Cobalt			Nickel			Vanadium		
	NSI base	NSI resample	G-BASE	NSI base	NSI resample	G-BASE	NSI base	NSI resample	G-BASE
Counts*	623	229	6411	623	229	6411	476	228	6411
Mean	10.6	10.3	19.4	26.3	27.8	23.5	39.0	45.6	88.2
SD	8.05	5.10	8.31	22.65	15.39	15.04	35.72	29.5	43.43
Skew†	7.29	0.79	0.98	10.28	0.99	6.14	6.87	2.2	3.85
Median	9.8	10.0	19.1	23.4	24.0	22.0	34.7	40.6	83.0
Q_n	5.76	4.92	7.54	13.06	13.3	11.10	22.52	22.6	33.3
Octile skew‡	0.05	0.01	0.0	0.13	0.33	0.07	-0.19	0.16	0.1
Analytical variance	5.5	5.5	0.8	3.5	3.5	0.3	9.45	9.45	0.2
Detection threshold	2.78	2.78	1.0	10.25	10.25	1.0	0.01	0.01	1.0

* There were 203 collocated observations of cobalt and nickel on both dates, and 155 of vanadium.

†Conventional index of skew based on third and second moments.

‡Octile skew defined in Equation (2).

Table 2 Cross-validation results for different auto-variogram models for soil cobalt concentration on each date.

Variogram estimator	Date 1*				Date 2†			
	Mean error	Median error	$\bar{\theta}$	$\tilde{\theta}^{\ddagger}$	Mean error	Median error	$\bar{\theta}$	$\tilde{\theta}^{\ddagger}$
Matheron	0.026	0.81	0.945	0.180	-0.006	0.47	1.26	0.331
$\widehat{\gamma}_{u,u}^{\mathrm{SRD}}$	0.007	0.81	2.70	**0.450**	-0.008	0.51	1.89	**0.488**
$\widehat{\gamma}_{u,u}^{\mathrm{MAD}}$	0.027	0.86	3.40	**0.540**	0.002	0.42	2.28	0.563
$\widehat{\gamma}_{u,u}^{Q_n}$	-0.021	0.78	9.68	1.624	0.003	0.32	1.87	**0.488**

* NSI baseline survey

† NSI resampling and G-BASE data, former rescaled to the same median value as the latter

‡ Under the null hypothesis the 95% confidence interval for $\tilde{\theta}$ (determined by bootstrapping) is 0.376–0.546 (Date 1) and 0.392–0.526 (Date 2). The latter limit is narrower because the data set is larger. Values in bold lie within this interval.

Table 3 Parameters of the linear models of coregionalization. The superscript 0 denotes the nugget component, and 1 the spatially dependent component with a the range of a spherical variogram and the distance parameter of an exponential. The superscripts denote date, 1 or 2.

Metal	$b^0_{1,1}$	$b^0_{2,2}$	$b^0_{2,1}$	Model type	a/metres	$b^1_{1,1}$	$b^1_{2,2}$	$b^1_{2,1}$
Nickel	72.2	40.8	0.01	Spherical	15145	123.8	98.9	110.6
Vanadium	250.6	237.8	244.1	Exponential	6005	318.1	1017.9	392.8
Cobalt	7.4	11.8	9.4	Exponential	7639	30.1	52.0	35.7

Figure captions

Figure 1 The Humber-Trent region of the G-BASE survey.

Figure 2 Empirical cumulative frequency distributions for (thin line) resampled NSI data and (thick line) G-BASE data on (top) Cobalt, (middle) Nickel and (bottom) Vanadium concentrations in soil of the Humber-Trent region.

Figure 3 Auto-variograms and fitted models for cobalt concentration in the soil. Top (one graph) on Date 1 (NSI Baseline) and, below (four separate graphs) on Date 2 (NSI resample + G-BASE).
Matheron's (1962) standard estimator: •; $\widehat{\gamma}_{u,u}^{\mathrm{SRD}}$: ∘; $\widehat{\gamma}_{u,u}^{\mathrm{P,MAD}}$: +; $\widehat{\gamma}_{u,u}^{\mathrm{P},Qn}$: ×.

Figure 4 Auto-variograms for Date 1 (Top), Date 2 (Middle) and Pseudo cross-variogram for both dates with fitted linear model of coregionalization. Soil Cobalt concentration.

Figure 5 Auto-variograms for Date 1 (Top), Date 2 (Middle) and Pseudo cross-variogram for both dates with fitted linear model of coregionalization. Soil Nickel concentration.

Figure 6 Auto-variograms for Date 1 (Top), Date 2 (Middle) and Pseudo cross-variogram for both dates with fitted linear model of coregionalization. Soil Vanadium concentration.

Figure 7 Auto-variograms change process, (solid line) calculated from the linear models of coregionalization and (symbol) estimated from collocated data. Top: Cobalt; Middle: Nickel; Bottom: Vanadium.

Figure 8 Maps of Cobalt concentration obtained by point co-kriging for (Top) Date 1

and (Bottom) Date 2. Northings and Easting are UK Ordnance Survey Grid (metres). Units of concentration are mg kg^{-1}.

Figure 9 Maps of change in Cobalt concentration obtained by point co-kriging for (Top) and kriging variance (Bottom). Note that the changes are rescaled so that zero on the map is equivalent to the mean change between the two dates for this region.

Figure 10 Maximum point cokriging variance of change in concentration plotted against sample effort. Solid line: equal division of total effort between two dates. Dotted line: 5-km square grid on Date 1 and sample sites on Date 2 including all Date 1 sites. Broken line: 5-km square grid on Date 1 and sample sites on Date 2 not coincident with Date 1 sites. Top: Cobalt. Bottom: Vanadium.

Figure 11 Maximum point cokriging variance of change in concentration plotted against sample effort for three hypothetical LMCR. Correlation of the spatially dependent components of variation are (top) 0.0, (middle) 0.4 and (bottom) 0.8. There are two sampling strategies, both with equal effort on both dates. The heavier line is for resampling of the date 1 locations on date 2, the finer line is for date 2 sample points offset from date 1