The offset correlation, a novel quality measure for planning geochemical surveys of the soil by 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

This paper presents a quality measure to plan geostatistical soil surveys when measures based on the kriging variance are not applicable. The criterion is the consistency of estimates made from two non-coincident instantiations of a proposed sample design. We consider square sample grids, one instantiation is offset from the second by half the grid spacing along the rows and along the columns. If a sample grid is coarse relative to the important scales of variation in the target property then the consistency of predictions from two instantiations is expected to be small, and can be increased by reducing the grid spacing. The measure of consistency is the correlation between estimates from the two instantiations of the sample grid, averaged over a grid cell. We call this the offset correlation, it can be calculated from the variogram. This quality measure is illustrated for some hypothetical examples, considering both ordinary kriging and factorial kriging of the variable of interest. The factorial kriging case is considered since, when planning a small-scale synoptic geochemical survey we may wish only to map components of the variation of the target variable at certain spatial scales. The quality measure is then computed for ordinary and factorial kriging with variograms estimated from data on nickel, chromium and cobalt content of soil in the north-east of England. Our results show how the offset correlation responds to sample density and the form of the variogram, and how larger correlations can be achieved for factorial kriging than ordinary kriging at a given density. The results for data on soil metals showed that an offset correlation of 0.8 could not be achieved (ordinary kriging) by sampling at 5-km intervals, the density at which all of England and Wales is sampled. However, if the objective were to map by factorial kriging the coarser-scale components of variation, driven primarily by parent material, then for two of the metals (Co and Cr) the 5-km grid was adequate, and the sample effort of the survey from which the data were taken (0.44 samples km^2) was excessive.

Keywords: Kriging, Quality measures, Sample size, Sample design.

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

1. Introduction

Geochemical survey of the soil entails the collection of soil samples for analysis, typically on a more-or-less uniform grid, and subsequent interpolation of the observed values to produce local predictions of the variables which are presented as a map. Since the seminal work of Burgess and Webster (1980) it is common to interpolate by kriging (e.g. Tao, 1995). Kriging is based on a linear model of the regionalization of the variable of interest (Goovaerts, 1997), of which a key component is the variogram model. The kriging prediction of a variable at an unsampled site is a linear combination of available data. The combination is found that minimizes the expected squared error of the prediction (the kriging variance), conditional on a variogram model of the variable (Webster and Oliver, 2007).

When a geochemical survey is planned it is necessary to make decisions about the sampling design. In particular it is necessary to select a sample density (e.g. Reimann, 2005). The total cost of processing and analysing the sampled material from a specified area depends on the sample density, as does the total cost of field work. Sample density also determines the quality of the resulting predictions. To make a rational choice of sample density we therefore require two things. First, we must know how some appropriate measure of quality of the final map improves with increased density. Second, we must be able to specify a value of that quality measure which represents an acceptable quality standard for the end user of the data.

In the case of geostatistical survey it is possible to compute *a priori* a relationship between map quality and sample density. If the variogram is known, perhaps from a reconnaissance survey or a previous study of a cognate landscape, the kriging variance at some unsampled site depends only on the spatial distribution of sample points around that site. One may therefore produce a graph of the kriging variance as a function of sample density. This approach to survey planning was proposed by McBratney et al. (1981) and

has been used subsequently (e.g. Di et al., 1989; Ruffo et al., 2005). The methodology has been extended to cover prediction by cokriging (McBratney and Webster, 1983), cases where the mean is not assumed to be stationary (Brus and Heuvelink, 2007) and where the variable is log-normal (Lark and Lapworth, 2012) and to account for uncertainty in the variogram model (Marchant and Lark, 2006, 2007; Zhu and Stein, 2006).

A graph of kriging variance against sample density is necessary but not sufficient for survey planning by this approach. It is also necessary to know what kriging variance is deemed acceptable by the end user of the data. Kerry et al. (2010) and Ruffo et al. (2005) provide examples from agriculture in which maximum acceptable standard errors for predictions of nutrient concentrations were specified then used to determine the maximum spacing of a sample grid which was consistent with this requirement. Black et al. (2008) describe a study in which a consortium of policy makers and regulators agreed what were acceptable standard errors for predictions of key soil quality indicators, and sample requirements were computed from a geostatistical model of available data. This is the general approach for sample design advocated by de Gruijter et al (2006) in which the data user identifies critical values of some quality measure for estimates from the sample, and the statistician identifies the sampling requirements to achieve this.

However, it is not always possible to express the quality requirements for a geochemical survey in terms of kriging variances or standard errors. This is for two general reasons. First, a geochemical survey is not, in general, undertaken for the benefit of a single end-user with clearly defined requirements in terms of information quality. Geochemical surveys, particularly at small scale, are typically undertaken to provide data which will serve a variety of purposes, not all forseen at the time of sampling. For example, the Geochemical Baseline Survey of the Environment (G-BASE), undertaken by the British Geological Survey in Great Britain, was initially planned to support geological mapping and mineral exploration (Johnson et al., 2005), but has subsequently proved invaluable for studies and applications on, *inter alia*, soil pollution (Breward, 2003), the nutritional

quality of crops grown on soil (Johnson et al., 2009) and forensic soil science (Rawlins and Cave, 2004). When the continuation of this survey was planned this was no longer done with a single end user or type of end user, in mind but with the awareness that the data set will constitute a general national capabability to tackle a variety of problems. It is unlikely that the diverse requirements of all end users, even if they could all be forseen at the time of survey planning, could be summarized in terms of a requisite kriging variance for the final kriged geochemical map.

Second, a geochemical survey may be planned to provide a synoptic overview of the geochemistry of a region, on the understanding that more intensive local surveys would be required for further specific applications such as the local evaluation of a resource or assessment of a local environmental risk. For example Reimann et al. (2007) reported a geochemical survey of the C-horizon of podzols in a 188 000-km^2 part of the Barents region (Russia and Finland). This area was sampled at a low density (1 sample per 300 km^2) to provide an overview of the variation of gold and palladium concentrations. The objective was to identify areas where more detailed investigation of these elements would be justified. In this context, as Reimann et al. (2007) state, the purpose of the survey was not to provide precise local estimates but rather to provide a map which represents geochemical patterns across a region at spatial scales of interest as a basis for planning further resource investigation in more intensive local surveys. It is not apparent that the quality requirement for the initial extensive survey could be stated in terms of a prediction error variance. Nonetheless, the utility of the resulting map will depend on sample density, and an appropriate quality measure is necessary to allow the selection of an operational sample density on rational grounds.

Smith and Reimann (2008) discussed the quality of geochemical surveys and proposed that the user is concerned with what they call the 'robustness' of the survey procedure. A procedure, a sampling design at some particular density, is robust if two surveys, conducted by the same procedure but at non-coincident sample locations, would pro-

duce maps which exhibit the same general pattern of geochemical variation. Smith and Reimann (2008) illustrate this idea by visual interpretation of geochemical maps produced at different densities. This concept has intuitive appeal. The scientist or other data user is aware that geochemical properties are spatially variable. His or her concern is to resolve an underlying pattern of variation, but one consequence of representing the geochemistry of a region with a finite sample is that some features of the variation are represented and others are missed. A useful measure of the quality of a sampling strategy is therefore the degree of consistency that could be expected between repeated surveys of the same region. This consistency will be small if the spacing between sample points is large relative to the scales at which the target property shows substantial variation, and can be improved by increasing the sample density. What is needed is a quality measure which reflects this idea of consistency, and which can be calculated as a function of sample density, given statistical information on the spatial variability of the variable of interest.

In this paper we propose such a statistical quality measure based on this concept of consistency. This measure is based on the idea of Smith and Reimann (2008) but is obtained from a variogram model of the target variable and refers to the consistency of maps produced by kriging. We suggest that this is a useful quality measure for circumstances, as described above, where it is not possible to express the data user's requirements in terms of a kriging variance. In particular it is an intuitively appealing measure of the quality of a survey procedure which may be communicated to data users who may have no experience of stating their requirements for the quality of estimates in terms of variances. The quality measure can be computed from reconnaissance data, or other information which allows a variogram of the target variable to be estimated or approximated. Where the variogram shows nested spatial structures the quality measure can be computed for maps of the longer-range structures, estimated from the data by factorial kriging (Goovaerts and Webster, 1994).

In the next section of this paper we develop the proposed quality measure and

examine its properties. We then illustrate it using geochemical data on the soil from the G-BASE survey of a part of eastern England.

2. Theory

2.1. The proposed quality measure

Consider a survey of a variable conducted on a square grid, Grid 1, of interval ξ. The kth node of grid 1 has coordinates $\mathbf{x}_{k,1}$. We propose that a measure of the consistency of this sampling design is the correlation that is expected between kriging predictions made from Grid 1, and predictions made from a second grid, Grid 2, which is a translation of Grid 1 by $\xi/2$ along the rows and the same distance along the columns so that its kth node has coordinates $\mathbf{x}_{k,2} = \mathbf{x}_{k,1} + \{\xi/2, \xi/2\}$.

Let $\mathbf{x}_0$ be a target location at which two kriged predictions of a variable are obtained. The first prediction, $\tilde{Z}_1(\mathbf{x}_0)$, is obtained by ordinary kriging from the n_1 nearest neighbouring observations on Grid 1, we denote this prediction subset of nodes of Grid 1 by the ordered set $X_{1,\mathbf{x}_0}$. We denote the $n_1 \times 1$ vector of ordinary kriging weights by $\boldsymbol{\lambda}_{1,\mathbf{x}_0}$. The lth element of $\boldsymbol{\lambda}_{1,\mathbf{x}_0}$ is the kriging weight applied to the observed value at the lth node in $X_{1,\mathbf{x}_0}$. The second prediction, $\tilde{Z}_{2,n_2}(\mathbf{x}_0)$, is obtained by ordinary kriging from the n_2 nearest neighbouring observations on Grid 2 with kriging weights in $\boldsymbol{\lambda}_{2,\mathbf{x}_0}$ which is $n_2 \times 1$. As for Grid 1, the prediction subset of nodes from Grid 2 is denoted $X_{2,\mathbf{x}_0}$.

Let $\mathbf{C}_{2,1,\mathbf{x}_0}$ denote a $n_2 \times n_1$ matrix of covariances such that $\mathbf{C}_{2,1,\mathbf{x}_0}\{i,j\}$ is the covariance between the observation at the ith node in $X_{2,\mathbf{x}_0}$. and the jth node in $X_{1,\mathbf{x}_0}$. Similarly let $\mathbf{C}_{1,\mathbf{x}_0}$ and $\mathbf{C}_{2,\mathbf{x}_0}$ denote the variance-covariance matrices of the observations in $X_{1,\mathbf{x}_0}$. and $X_{2,\mathbf{x}_0}$ respectively. These matrices can be populated directly given the coordinates of the grid points and a (second-order stationary) variogram function for the variable of interest.

Given the notation above, the variances of $\tilde{Z}_1(\mathbf{x}_0)$ and $\tilde{Z}_2(\mathbf{x}_0)$ can be computed as

$$\sigma^2_{\tilde{Z}_1}(\mathbf{x}_0) \quad = \quad \boldsymbol{\lambda}^{\mathrm{T}}_{1,\mathbf{x}_0}\mathbf{C}_{1,\mathbf{x}_0}\boldsymbol{\lambda}_{1,\mathbf{x}_0}, \text{ and}$$

$$\sigma^2_{\tilde{Z}_2}(\mathbf{x}_0) \quad = \quad \boldsymbol{\lambda}^{\mathrm{T}}_{2,\mathbf{x}_0}\mathbf{C}_{2,\mathbf{x}_0}\boldsymbol{\lambda}_{2,\mathbf{x}_0}, \tag{1}$$

and the covariance of $\tilde{Z}_1(\mathbf{x}_0)$ and $\tilde{Z}_2(\mathbf{x}_0)$ is

$$C_{\tilde{Z}_1,\tilde{Z}_2}(\mathbf{x}_0) \; = \; \boldsymbol{\lambda}^{\mathrm{T}}_{2,\mathbf{x}_0}\mathbf{C}_{2,1,\mathbf{x}_0}\boldsymbol{\lambda}_{1,\mathbf{x}_0}. \tag{2}$$

The correlation of the two kriging predictions may then be obtained as

$$\rho_{\tilde{Z}_1,\tilde{Z}_2}(\mathbf{x}_0) \; = \; \frac{C_{\tilde{Z}_1,\tilde{Z}_2}(\mathbf{x}_0)}{\sqrt{\sigma_{\tilde{Z}_1}\sigma_{\tilde{Z}_2}}}. \tag{3}$$

In Figure 1 we show a map of the correlation of kriged estimates from two grids, each of interval 50 units, one grid translated from the other by 25 units along the rows and the same distance along the columns. The correlations are mapped at locations in a cell of one of those grids, with one node of the second grid at the centre. At all locations in the figure the mapped correlation is between the prediction by ordinary kriging from the nearest 16 nodes in the first grid and the nearest 16 nodes in the second grid, each set of 16 nodes being a regular 4×4 array. This is for a hypothetical example in which the variogram of the variable is an isotropic spherical function. The linear model of regionalization for such variables comprises two independent additive components. The first, the nugget component with variance c_0 is spatially uncorrelated over the shortest distances between observations. The second spatially correlated component, with variance c_1 shows spatial dependence over distances up to the range, a. The overall variance of the variable (the *a priori* variance) is $c_0 + c_1$. At longer distances than the range observations of the variable are not spatially dependent. The variogram function is

$$\gamma(h) \quad = \quad c_0 + c_1 \mathrm{Sph}\left(h|a\right), \tag{4}$$

where

$$\begin{aligned} \mathrm{Sph}(h|a) \quad &= \quad \left\{\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right\}, \qquad h \leq a, \\ &= \quad 1, \qquad\qquad\qquad\qquad h > a. \end{aligned} \tag{5}$$

In the example here $a \; = \; 150$ units, $c_0 = 0.2$ and $c_1 = 0.8$. Figure 1 shows the variation of $\rho_{\tilde{Z}_1,\tilde{Z}_2}(\mathbf{x}_0)$ across the grid cell. Note that the correlation decreases as one approaches

a node of either prediction grid, and is largest between the nodes, where the influence of the two grids is most similar.

In this paper our proposed quality measure for a survey on a regular grid of interval ξ is the average value of the correlation $\rho_{\tilde{Z}_1,\tilde{Z}_2}(\mathbf{x}_0)$ across a cell of one grid, where the two grids are of interval ξ and one is a translation of the other by $\xi/2$, as described above. In this paper we compute the correlation for kriging predictions from the nearest 4×4 subset of nodes in each array. We call this measure the offset correlation.

2.2. Hypothetical examples

In Figure 2a offset correlations are plotted for grids of different spacing for kriging predictions of a regionalized variable with a spherical variogram, as defined in Equation (4). The variogram parameters are $a = 100$ units, and c_1 varies from 1.0 to 0.1, with $c_0 = 1.0 - c_1$. As expected the offset correlation declines with increasing grid spacing, for a given variogram, and also declines as the nugget variance c_0 increases relative to the correlated variance c_1. Note that the offset correlation goes to zero when $\xi = a\sqrt{2}$. Some elementary geometry shows that for this grid the distance between any node in subset $X_{1,\mathbf{x}_0}$ and the nearest node in subset $X_{2,\mathbf{x}_0}$ is a, and so, for this or any coarser grid, the covariance between any two observations on Grid 1 and Grid 2 is zero. When the nugget variance is zero then an offset correlation of 0.8 can be achieved with a 50-unit square grid. However, when the nugget variance is half of the *a priori* variance the grid interval must be about 22 units to achieve the same offset correlation.

Figure 2b shows comparable plots for predictions of a variable with an exponential variogram

$$\gamma(h) \quad = \quad c_0 + c_1 \left(1 - \exp\left\{-h/r\right\}\right), \tag{6}$$

with r=30 units and the same range of values for c_0 and c_1 as for the examples with a spherical variogram. The effective range of this variogram (at which $\gamma(h) \approx 0.95(c_0 + c_1)$) is 90 m. The behaviour of the offset correlation is similar to Figure 2b. The main difference

is that, since the covariance of a process with an exponential variogram declines to zero asymptotically, the offset correlation does not go exactly to zero above some grid spacing.

Figure 2c shows plots of the offset correlation for random variables with a double-spherical variogram. The double-spherical variogram describes a linear model of regionalization which comprises three mutually independent additive components, a nugget component and two components, with variance c_1 and c_2, which are spatially correlated at different scales with range parameters a_1 and a_2 respectively. The double spherical variogram model is

$$\gamma(h) \quad = \quad c_0 + c_1 \mathrm{Sph}\,(h|a_1) + c_2 \mathrm{Sph}\,(h|a_2) \tag{7}$$

In this example we considered variables with a_1=50 units, a_2=125 units, c_0 =0.1 units, and various values of c_1 and c_2 such that the *a priori* variance is 1.0 in all cases.

Consider a situation in which the longer-range component of a variable with a double-spherical variogram represents the source of variation of primary interest. For example, it might represent variation due to geochemical differences between types of parent material, whereas the shorter-range component represents effects of diffuse pollution. If our primary concern is to map the coarser-scale pattern, then this can be done by kriging analysis, or factorial kriging (e.g. Goovaerts, 1997). Goovaerts and Webster (1994) used factorial kriging to estimate separate components of geochemical variation in the soil of south-east Scotland. The component of the linear model with a shorter range was interpreted as a land-management effect, and the longer-range component as a geological effect. In some contexts we are interested in the former, but not the latter, such as when soil geochemistry is mapped as a surrogate for investigation of the geochemical variation between parent materials. In such circumstances the quality measure of interest for the geochemical survey is the offset correlation between the factorial kriging estimates of the component of interest. This can be calculated by substituting the vectors of factorial kriging weights $\boldsymbol{\lambda}^2_{1,\mathbf{x}_0}$ and $\boldsymbol{\lambda}^2_{2,\mathbf{x}_0}$ into Equations (1) and (2), where the superscript is an index not a power, and $\boldsymbol{\lambda}^g_{1,\mathbf{x}_0}$ is the factorial kriging weight to estimate the gth component of a

nested random variable at $\mathbf{x}_0$ from observations on Grid 1. The factorial kriging weights are obtained by solving the factorial kriging equations, as described by Goovaerts (1997) and Webster and Oliver (2007).

Figure 2d shows the offset correlations for factorial kriging predictions corresponding to the ordinary kriging predictions in Figure 2c. In all cases the factorial kriging prediction is for the coarsest-scale component, with a range of 125 units.

To achieve an offset correlation of 0.8 for ordinary kriging predictions of the double spherical random variable, with c_2=0.7, requires a grid of interval 40 units. To achieve the same standard for factorial kriging predictions of the coarsest scale component requires a grid interval of just under 50 units. Figures 2c and 2d show that the offset correlation decays less rapidly with grid interval for the factorial kriging case, other factors being equal. Note, however, that the offset correlation at the finest grid spacing is more sensitive to the relative values of c_2 and c_1 than is the offset correlation for the ordinary kriging predictions.

This section has introduced the offset correlation, and examined its behaviour for some hypothetical examples, considering both ordinary kriging and factorial kriging analysis to estimate scale-specific components of a variable. In the next section we examine a case study with data from a geochemical survey of the soil, and use validated variograms of soil properties to examine the offset correlations for predictions by ordinary and factorial kriging.

3. A case study with soil data

3.1. The soil data

We used soil data from the British Geological Survey's G-BASE survey of the Humber-Trent region, approximately 15 800 km^2 (North East England). A more detailed account of this G-BASE survey is given by Rawlins et al (2003) and the G-BASE procedures are described by Johnson et al. (2005). Alternate 2-km squares of the UK Ordnance

Survey grid were sampled at a single site within the square. At each sample site five soil cores were collected from the centre and corners of a 20-m square. Each core was 15-cm long, excluding surface litter. The five cores at each site were bulked, and this material was subsequently air-dried, disaggregated and sieved to pass 2 mm. and sub-sampled by coning and quartering. A 50-g sub-sample was ground in an agate planetary ball mill until 95% of the material was finer than 53 μm. Concentrations (totals) of 26 major and trace elements were determined for each sample by wavelength dispersive X-Ray Fluorescence Spectrometry. We used data from 5892 sites.

3.2. Statistical analysis

3.2.1. Exploratory analysis and variogram estimation, modelling and validation. For purposes of this paper we present analyses of data on the concentrations of chromium, cobalt and nickel. Summary statistics for these data are presented in Table 1. The summary statistics include the octile skew (Brys et al., 2003) which is a measure of the symmetry of the 1st and 7th octiles of the data about the median. The octile skew is a robust measurements of skewness, which is insensitive to outlying observations but measures rather the degree of asymetry of the underlying distribution. Data are considered for transformation if the conventional coefficient of skewness lies outside the interval [−1,1], (Webster and Oliver, 2007). Lark et al (2006) found that a corresponding interval for the octile skew is [−0.2,0.2]. If the coefficient of skewness for a variable is outside the interval [−1,1] but the octile skew is small then this suggests that the data have an underlying distribution that is more-or-less symmetrical but that there are outliers present. The three variables considered here all have small octile skew, with absolute values less than 0.1, which suggests that a transformation is not appropriate for the data. However, the conventional coefficients of skewness are large for nickel and, particularly, for chromium, which suggests that these observations may include outlying values, perhaps from point pollution.

Exploratory geostatistical analysis suggested that these data do not show pronounced anisotropy, and so we estimated isotropic variograms using the conventional

method of moments estimator due to Matheron (1962) as well as three robust estimators, proposed by Cressie and Hawkins (1980), Dowd (1984) and Genton (1998). Robust estimators were considered because of the suggestion from the exploratory analysis that the data may contain outliers.

Variogram models for each experimental variogram were selected on the basis of the Akaike Information Criterion (Webster and Oliver, 2007). Double spherical variogram models were selected in all cases, and fitted to the estimated variograms by weighted least squares with the FVARIOGRAM procedure in GenStat (Payne, 2010). The variogram models were then cross-validated. The XVOK2D program in the GSLIB library (Deutsch and Journel, 1992) was used for this purpose. The standardized square cross validation prediction error, $\theta(\mathbf{x})$ was computed from the cross-validation prediction , $\tilde{Y}(\mathbf{x})$, of each observation in the data set, $Y(\mathbf{x})$, and the corresponding kriging variance $\sigma^2_{\mathrm{K}}(\mathbf{x})$.

$$\theta(\mathbf{x}) \;=\; \frac{\left(\tilde{Y}(\mathbf{x}) - Y(\mathbf{x})\right)^2}{\sigma^2_{\mathrm{K}}(\mathbf{x})}, \tag{8}$$

We computed normal Q-Q plots of the cross-validation errors (Figure 3). These indicated that the prediction errors appeared to be normally distributed, although with some effects of outliers. Lark (2000a) showed that the median value of the standardized squared prediction error is the most appropriate diagnostic to evaluate a variogram from cross-validation output, and when the prediction error are predominantly normal the expected value of this statistic is 0.455. The cross-validation results were used to select a variogram model from among the set of those fitted to the experimental variograms obtained by the different estimators. The variogram model thus selected was then used to compute the offset correlations for each variable, as described below.

3.2.2. Offset correlations. Offset correlations were computed for square grids with intervals from 500 m to 30 km. The average offset correlation was computed across a cell of one of the grids, as described in section 2.1. Offset correlations were computed both for ordinary kriging predictions, and for factorial kriging predictions of the component of the linear model of regionalization with the longest range.

3.3. Results

Table 2 shows the cross-validation results for all three variables. In all cases the variogram model based on the estimator of Cressie and Hawkins (1980) was selected because the median standardized squared prediction error was closest to 0.455. The selected model and the associated point estimates, as well as the estimates by Matheron's estimator, are shown in Figure 4, and the model parameters are presented in Table 3. The difference between the models can be attributed to outlying data which have a larger effect on Matheron's estimator than on the robust estimator. The value of the median standardized squared prediction error for kriging with the selected variogram model based on a robust estimator suggests that this gives a reliable account of the uncertainty of the kriging predictions.

The offset correlations are plotted against sample density in Figure 5. Figure 5a shows the offset correlations for ordinary kriging, and Figure 5b shows the offset correlations for factorial kriging of the longest-range component. Two sample densities are indicated on these graphs. One is 0.04 samples km^{-2}, the sample density of the National Soil Inventory in England and Wales (McGrath and Loveland, 1992). The second is $0.4\dot{4}$ samples km^{-2}, the sample density for soils in the G-BASE survey (Johnson et al., 2005).

If we regard an offset correlation of 0.8 as a standard for selecting a sample density for ordinary kriging then it is clear from Figure 5a that the sample density of the NSI is not adequate to meet this standard for all three elements. The offset correlations are 0.74 for chromium and nickel and 0.79 for cobalt. However, the standard is easily met with the G-BASE sample density, the offset correlations are 0.89, 0.91 and 0.93 for nickel, chromium and cobalt respectively.

The graph in Figure 5a shows that nickel is the most challenging of these three elements, in that it has the smallest offset correlation at any given sample density. However, if the sample density were reduced to 0.12 samples km^{-2}, a reduction of sample effort by a factor of nearly 4 relative to the G-BASE survey, then the offset correlation standard of

0.8 would be achieved for nickel. This could be useful information when planning a survey on a neighbouring region, or over similar parent materials.

If we are concerned only to predict and map, by factorial kriging, the broader-scale variations of the elements, represented by the longer-range component in the linear model of regionalization (a range of 18–20 km) then the offset correlations of interest are those in Figure 5b. This shows that the NSI sample density is adequate to meet the 0.8 offset correlation standard for cobalt and chromium, but not for nickel. The offset correlations at G-BASE sample density are large (0.95–0.97), and an offset correlation of 0.8 could be achieved for all three elements by sampling at 0.049 samples km^{-2} , a nine-fold reduction in sample effort relative to G-BASE. Note that there is very little increase in the offset correlation for the factorial kriging estimates when the sample density is larger than the G-BASE density.

4. Discussion

The offset correlation is a proposed measure for the quality of a geostatistical survey. It can be computed for a proposed sample scheme given only the variogram of the variable of interest. It is an intuitively appealing measure of the extent to which a survey can be expected to provide a map of spatial variation that is robust to arbitrary differences between realizations of a particular sample design. Users of data are familiar with the concept of correlation, and its measurement on an interval [0,1], and so it is proposed that this measure could be useful for discussing the sampling requirements for a survey with scientists or other data users with little or no statistical background. In particular it could be useful in circumstances where it is difficult for the data user to express their requirements for information quality in terms of standard errors of predictions.

In this paper we considered simple grid surveys, but the same approach could be used to evaluate alternative sample designs such as unaligned sampling in which the good spatial coverage required for local prediction is combined with an element of randomization. Offset correlations could be computed between pairs of realizations of this sample design.

In this paper we have considered the variogram parameters as fixed but unknown quantities to be estimated, and we have taken no account of parameter uncertainty. Given the large sample available this was not unreasonable. In circumstances where the variogram has been estimated from a smaller reconnaissance sample we should try to account for parameter uncertainty. This may be done most conveniently in a Bayesian framework in which the variogram parameters are treated as random variables. Bayesian estimation allows us to obtain a posterior distribution of the variogram parameters (e.g. Orton et al, 2007; Minasny et al., 2011), and a corresponding distribution of the offset correlation could be computed by sampling this distribution. It would also be possible to make general recommendations about the sampling effort required to achieve a particular offset correlation on the basis of average variograms culled from the literature (McBratney and Pringle, 1999), fuzzifications of the variogram (Lark, 2000b) or from variograms of ancillary variables such as airborne gamma radiometry which we might reasonably treat as a proxy for the spatial variation of soil geochemistry (Rawlins et al, 2007).

5. Conclusions

We have derived the offset correlation, a statistical measure of the robustness of geostatistical prediction to arbitrary variations between realizations of a sample design (here a regular grid). We have illustrated how this measure behaves from hypothetical examples and a real case study on soil geochemistry. The offset correlation can be computed from the variogram for the target variable for either ordinary kriging or for prediction by factorial kriging of a specific component of the linear model of regionalization that is of interest. As expected, a comparison between the offset correlations for factorial and ordinary kriging shows that a coarser sample grid can be used to map the broad-scale components of a variable than is needed to achieve the same offset correlation for all components. It is proposed that this could be a useful quality measure on which to base the planning of a geostatistical survey in cases where it is difficult or impossible for the end user of the information to frame their quality requirements in terms of standard errors or

variograms of the prediction error.

Acknowledgements

The contribution of all British Geological Survey staff involved in the geochemical sampling of the Humber Trent region, and the subsequent soil analysis, is acknowledged. This paper is published with the permission of the Director of the British Geological Survey (NERC).

References

Black, H., Bellamy, P., Creamer, R., Elston, D., Emmett, B., Frogbrook, Z., Hudson, G., Jordan, C., Lark, M., Lilly, A., Marchant, B., Plum, S., Potts, J., Reynolds, B., Thompson, P., Booth., P. 2008. Design and operation of a UK soil monitoring network. Environment Agency, Bristol. Science Report – SC060073.

Breward, N. 2003. Heavy-metal contaminated soils associated with drained fenland in Lancashire revealed by BGS Geochemical Survey. Applied Geochemistry 18, 1663–1670

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

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.

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

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

Deutsch, C.V, Journel, A.G. 1992. 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.

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.

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

Goovaerts, P., Webster, R. 1994. Scale-dependent correlation between topsoil copper and cobalt concentrations in Scotland. European Journal of Soil Science 45, 79–95.

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.

Johnson, C. C.; Fordyce, F. M.; Rayman, M.P. 2009 Symposium on 'Geographical and geological influences on nutrition' : factors controlling the distribution of selenium in the environment and their impact on health and nutrition. Proceedings of the Nutrition Society 69. 119-132

Lark, R.M. 2000a. 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. 2000b. Designing sampling grids from imprecise information on soil variability: an approach based on fuzzy kriging variance. Geoderma 98, 35–59

Lark, R.M., Lapworth, D.J. 2012. Quality measures for soil surveys by lognormal kriging. Geoderma 173-174, 231–240.

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.

Kerry, R., Oliver, M.A., Frogbrook, Z.L. 2010. Sampling in precision agriculture. In: Geostatistical Applications for Precision Agriculture (ed M.A. Oliver), pp. 35–63. Springer, Dordrecht.

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.

McBratney, A. B., Pringle, M. J. 1999. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. Precision Agriculture 1, 125-152.

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

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.

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

Minasny, B., Vrugt, J.A., McBratney, A.B., 2011. Confronting uncertainty in model-based geostatistics using Markov Chain Monte Carlo simulation. Geoderma 163, 150–162.

Orton, T.G., Rawlins, B.G., Lark, R.M. 2009. Using measurements close to a detection limit in a geostatistical case study to predict selenium concentration in topsoil. Geoderma 152, 269–282

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

Rawlins, B.G., Cave, M. 2004. Investigating multi-element soil geochemical signatures and their potential for use in forensic studies. In: (eds) Pye, K., Croft, D. Forensic Geoscience Principles, Techniques and Applications. Geological Society, London. Special Publications, 232, 197–206.

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.

Rawlins, B.G., Lark, R.M., Webster, R. 2007. Understanding airborne radiometric survey signals across part of eastern England. Earth Surface Processes and Landforms 32, 1503–1515.

Reimann, C. 2005. Sub-continental scale geochemical mapping: sampling, quality control and data analysis issues. Geochemistry: Exploration, Environment, Analysis 5, 311–323.

Reimann, C., Melezhik, V., Niskavaara, H. 2007. Low-density regional geochemical mapping of gold and palladium highlighting the exploration potential of northernmost Europe. Economic Geology 102, 327–334.

Ruffo, M.L., Bollero, G.A., Hoeft, R.G., Bullock, D.G. 2005. Spatial variability of the Illinois soil nitrogen test: implications for soil sampling. Agronomy Journal 97, 1485–1492.

Smith, D.B., Reimann, C. 2008. Low-density geochemical mapping and the robustness of geochemical patterns. Geochemistry: Exploration, Environment, Analysis 8, 219–227.

Tao, S. 1995. Kriging and mapping of copper, lead and mercury contents in surface soil in Shenzhen area. Water, Air and Soil Pollution 83, 161–172.

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

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 on soil data.

	Cr	Co	Ni
	mg kg^{-1}		
Mean	75.03	19.52	23.72
Median	72.00	19.12	22.00
SD	54.06	8.26	14.17
Skewness	28.23	0.91	3.01
Octile skew	0.02	0.00	0.07

Table 2. 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
Cr	4.45	0.24	8.24	0.44	10.78	0.57	9.41	0.51
Co	1.01	0.32	1.6	0.48	1.88	0.58	1.98	0.63
Ni	1.15	0.23	2.42	0.46	3.7	0.69	3.41	0.64

Table 3. Parameters of the selected variogram model for each element.

	Cr	Co	Ni
Estimator	Cressie-Hawkins		
Model type	Double spherical		
c_0	199.5	12.9	11.6
c_1	176.9	12.3	42.5
c_2	378.3	35.4	82.7
a_1	1 813	4 332	2 535
a_2	21 409	21 228	16 115

Figure captions

1. Correlations across four unit cells of a square grid of length 50 units (grid nodes indicated by a $\times$) between ordinary kriging estimates of a variable obtained from the nearest 16 nodes of the grid, and estimates derived from the nearest 16 nodes of a second grid with the same interval but translated along the rows and columns by 25 units (grid nodes indicated by a $+$). The variable has a spherical variogram $\gamma(h) = 0.2 + 0.8\,\mathrm{Sph}(h|125)$ where $\mathrm{Sph}(\cdot|\cdot)$ is defined in Equation (5).

2a. Average correlation across a unit cell of a grid between ordinary kriging predictions of a variable from the grid and those from a grid translated by half the grid interval along the rows and the same distance along the columns (offset correlations). The average correlation is plotted as a function of grid interval. Example for a variable with a spherical variogram with a range parameter of 100 units. Results are given for variograms with different values of the correlated variance, c_1 shown by different symbols. In all cases the *a priori* variance of $c_0 + c_1 = 1$.

2b. Offset correlations as in Figure 2a, but for a variable with an exponential variogram with a distance parameter of 30 units. Results are given for variograms with different values of the correlated variance, c_1 shown by different symbols. In all cases the *a priori* variance of $c_0 + c_1 = 1$.

2c. Offset correlations as in Figure 2a, but for a variable with a double spherical variogram with ranges 50 and 125 units and nugget variance $c_0 = 0.1$. In all cases $c_1 + c_2 = 0.9$ so the *a priori* variance is 1. Different values of c_2 are indicated by symbols in the plot.

2d. Offset correlations for a variable with a double spherical variogram, as in Figure 2c, but these are correlations for the factorial kriging predictions of the component with a range 125 units.

3. Empirical normal QQ plots for cross-validation errors of chromium, cobalt and nickel. In each case the empirical quantile of a datum is plotted against the corresponding normal quantile of a random variable with mean and standard deviation equal to robust estimates of these parameters from the data.

4. Variogram estimates for chromium, cobalt and nickel. The solid discs show estimates by Matheron's estimator. The open circles are estimates obtained by the robust

estimator selected from the cross-validation statistics (Cressie and Hawkins, 1980; in all cases). The model fitted to the robust estimates is also shown.

5a. Offset correlations for ordinary kriging estimates of chromium, cobalt and nickel in the soils of the Humber-Trent region plotted against the sampling density of a square grid. The densities of the G-BASE survey and the National Soil Inventory of England and Wales are indicated by vertical lines.

5b. Offset correlations for factorial kriging estimates of the long-range (18–20 km) component of the linear models of regionalization for chromium, cobalt and nickel plotted against the sampling density of a square grid.

Figure 1

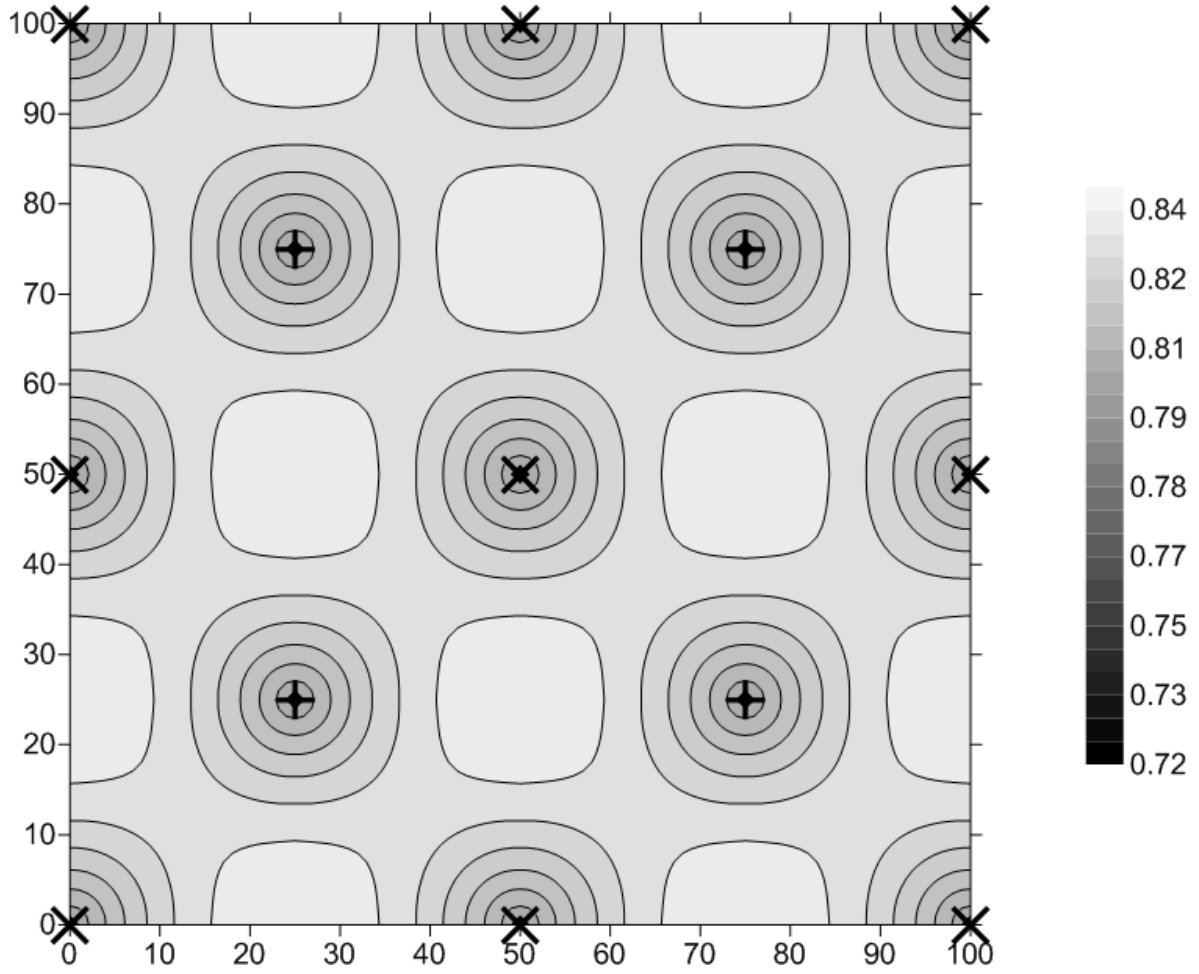

100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
90
100
0.84
0.82
0.81
0.79
0.78
0.77
0.75
0.73
0.72

Figure 2a

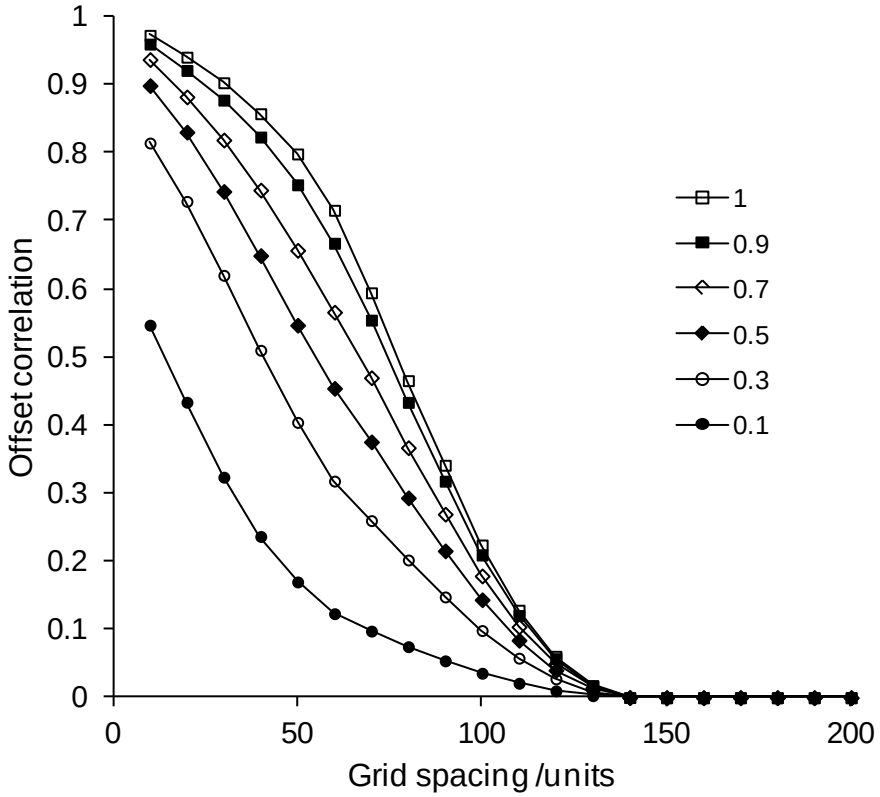

Offset correlation
Grid spacing /units
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
1
0.9
0.7
0.5
0.3
0.1

Figure 2b

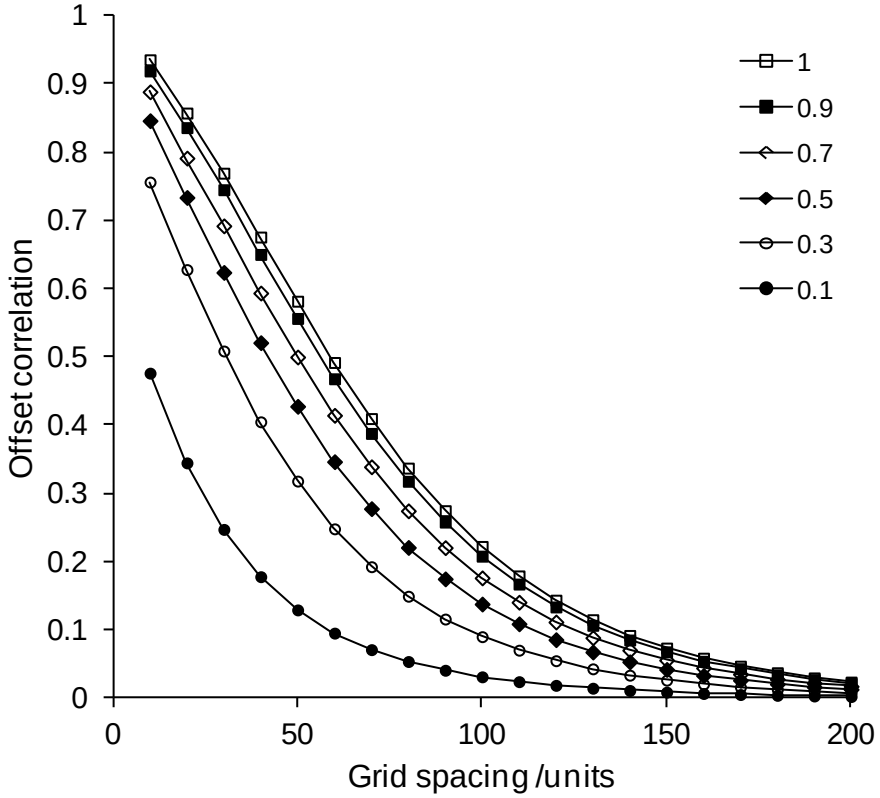

Offset correlation
Grid spacing /units
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
1
0.9
0.7
0.5
0.3
0.1

Figure 2c

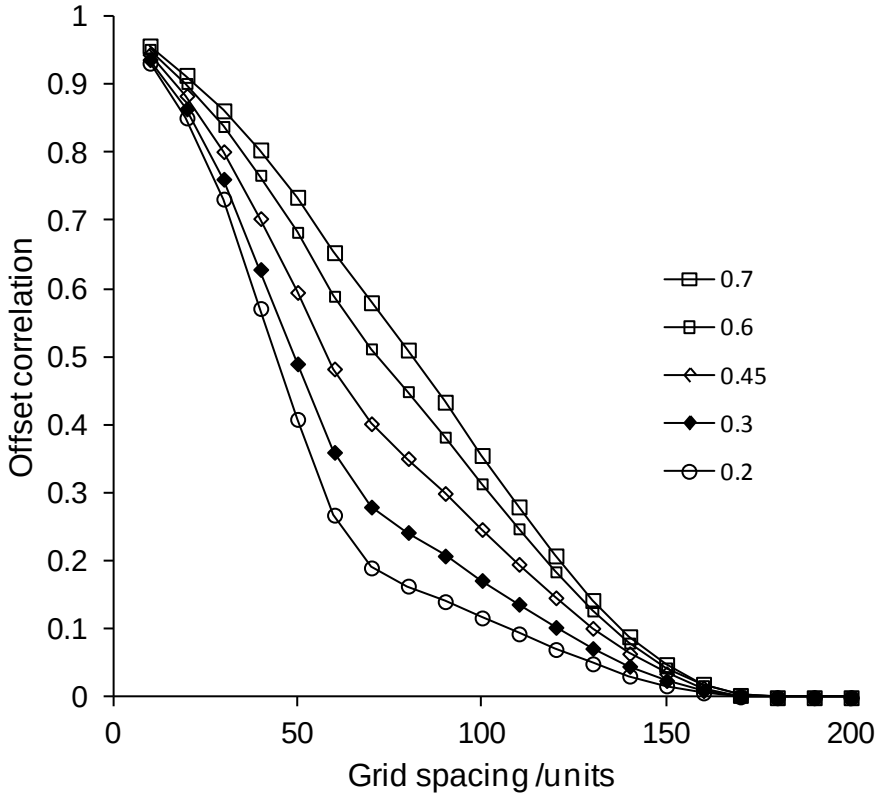

Offset correlation
Grid spacing /units
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
0.7
0.6
0.45
0.3
0.2

Figure 2d

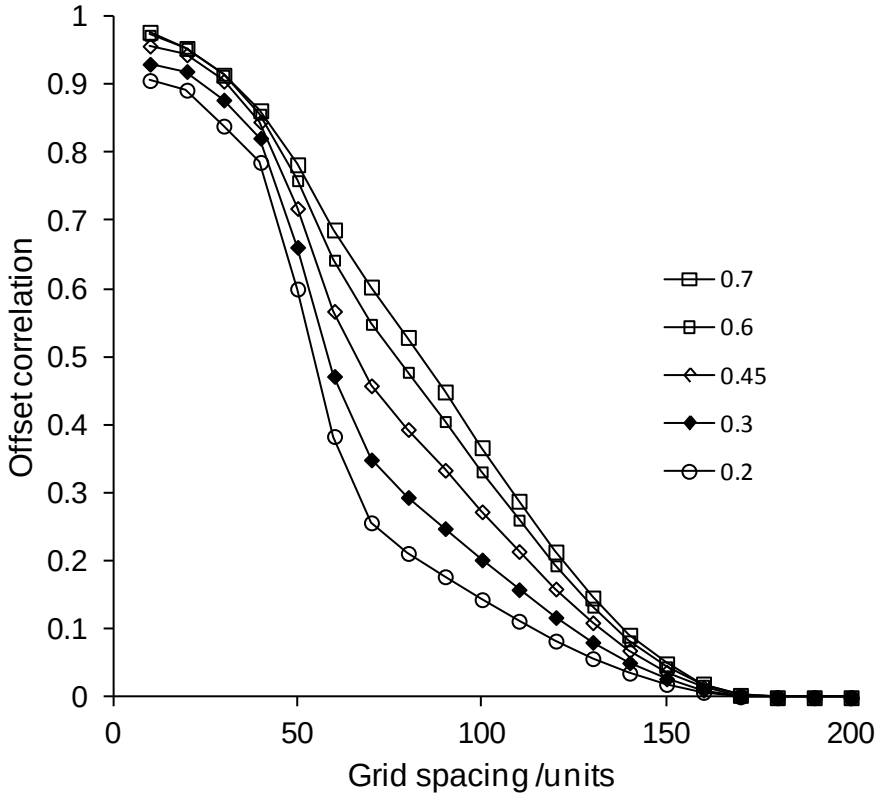
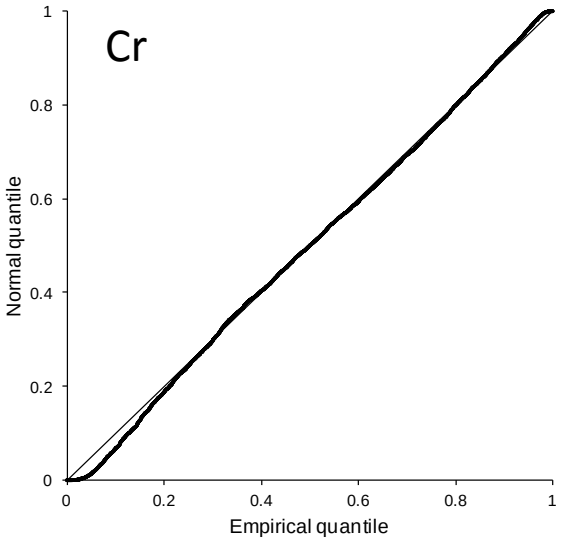
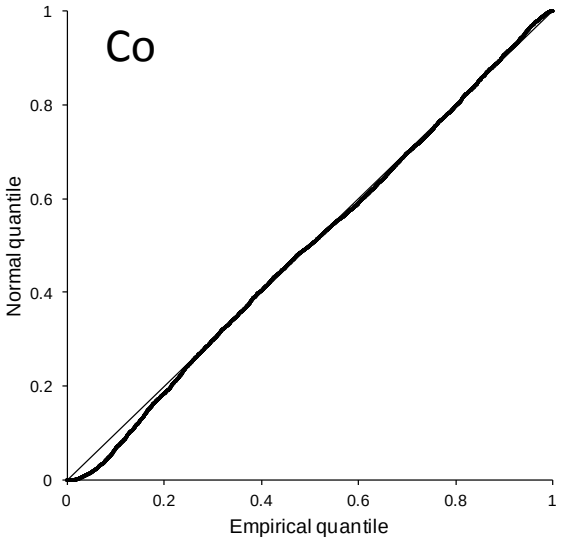

Offset correlation
Grid spacing /units
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
0.7
0.6
0.45
0.3
0.2

Figure 3


Cr
Normal quantile
Empirical quantile
0
0.2
0.4
0.6
0.8
1


Co
Normal quantile
Empirical quantile
0
0.2
0.4
0.6
0.8
1

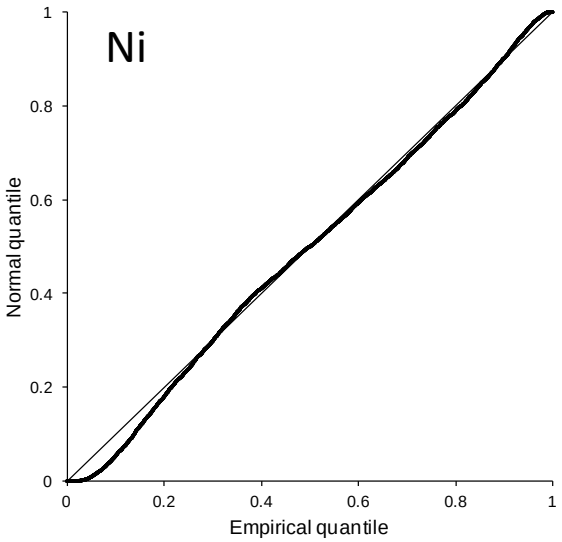
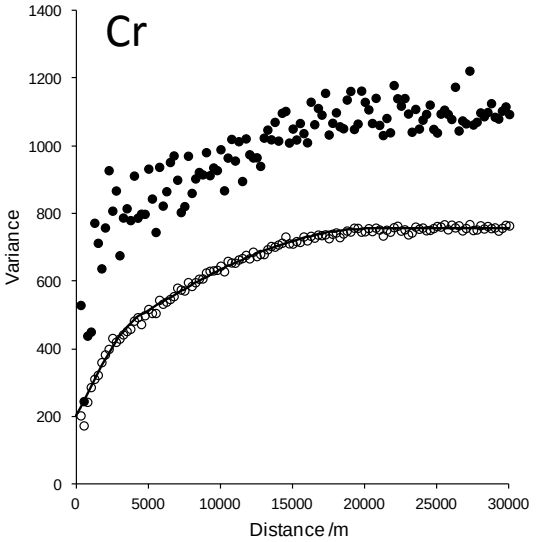
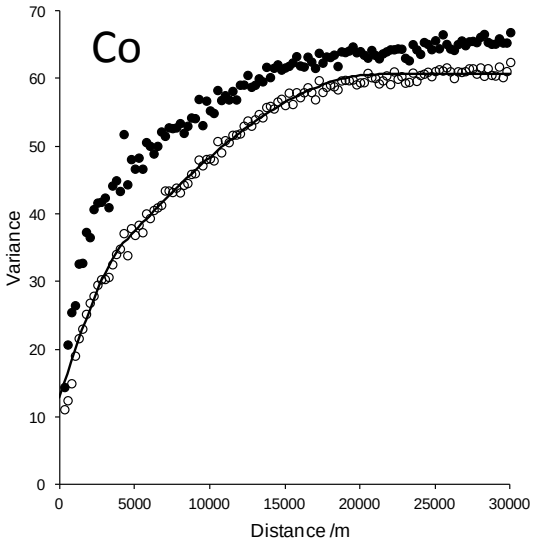

Ni
Normal quantile
Empirical quantile
0
0.2
0.4
0.6
0.8
1

Figure 4


Cr
Variance
Distance /m
0
200
400
600
800
1000
1200
1400
0
5000
10000
15000
20000
25000
30000


Co
Variance
Distance /m
0
10
20
30
40
50
60
70
0
5000
10000
15000
20000
25000
30000

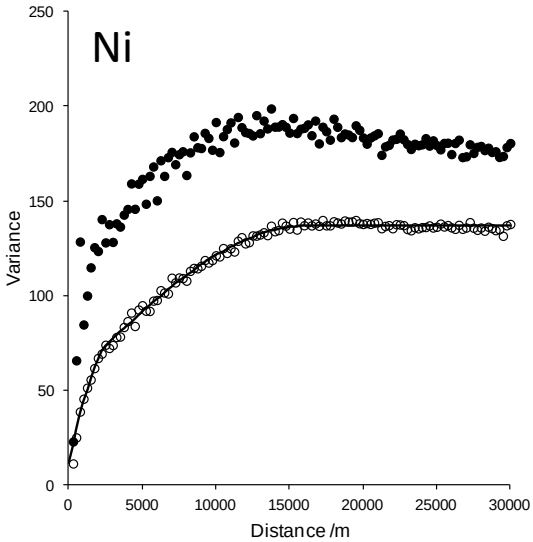

Ni
Variance
Distance /m
0
50
100
150
200
250
0
5000
10000
15000
20000
25000
30000

Figure 5a

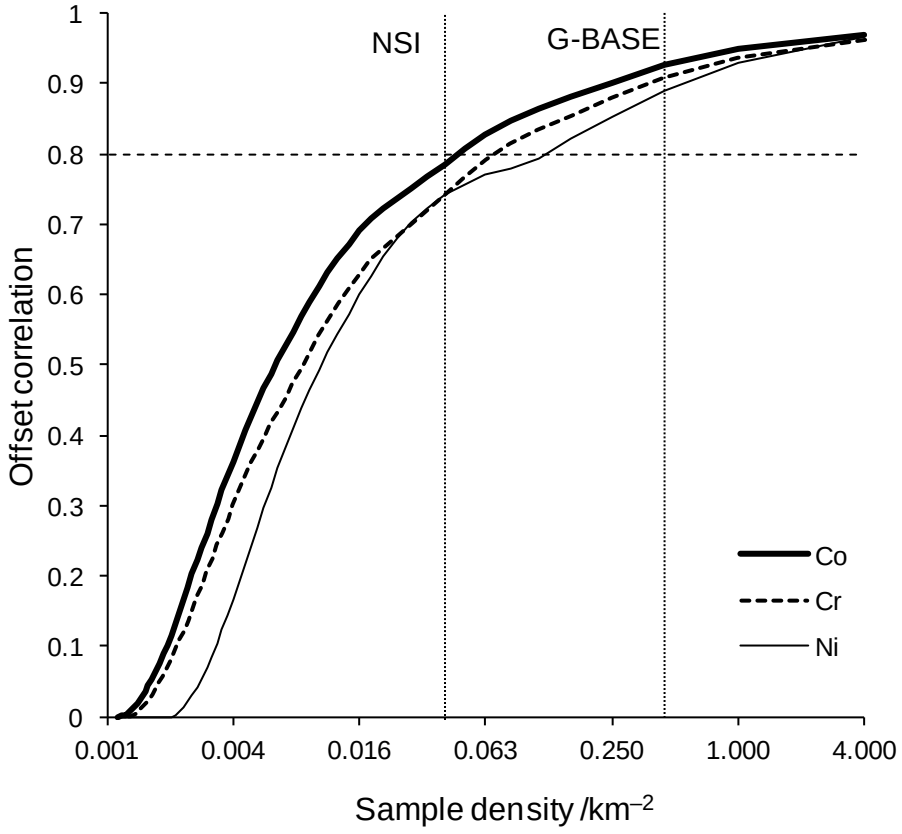
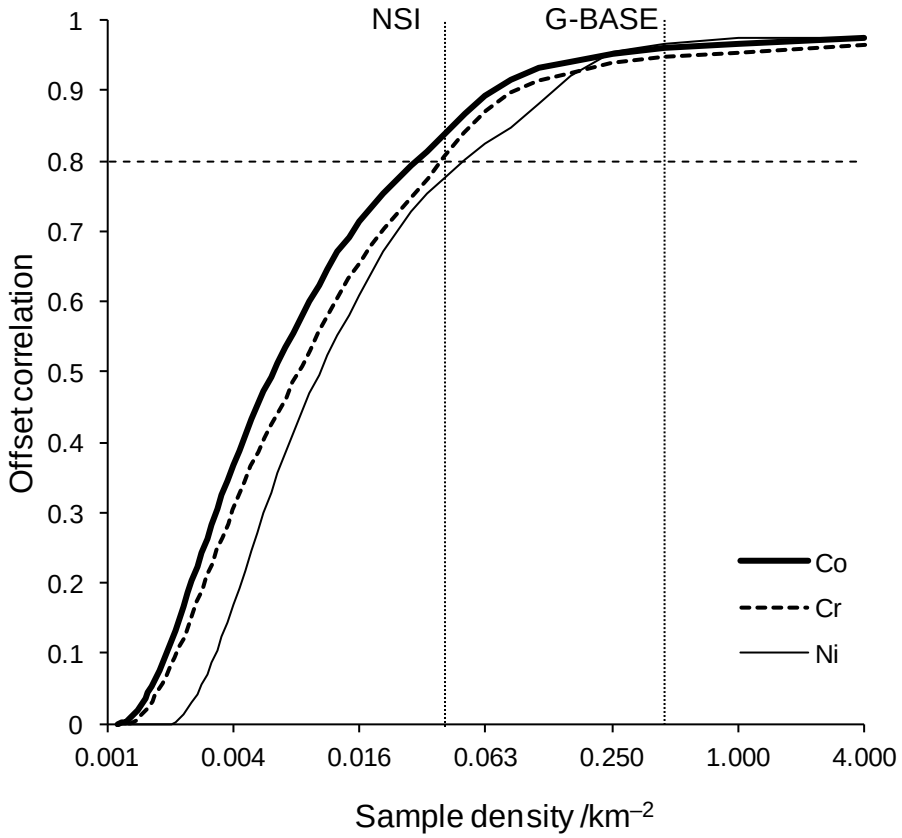

NSI
G-BASE
Offset correlation
Sample density /km^{-2}
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.001
0.004
0.016
0.063
0.250
1.000
4.000
Co
Cr
Ni

Figure 5b


NSI
G-BASE
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Offset correlation
0.001
0.004
0.016
0.063
0.250
1.000
4.000
Sample density /km−2
Co
Cr
Ni