QUALITY ASSURANCE AND QUALITY CONTROL (QA/QC) FOR IN-SITU GEOCHEMICAL METHODS, ESTIMATION OF MEASUREMENT UNCERTAINTY AND CONSTRUCTION OF PROBABILITY RISK ASSESSMENT MAPS

Alecos Demetriades and Helen Karamanos
Institute of Geology and Mineral Exploration, Greece

ABSTRACT
This document gives guidelines for taking duplicate measurements by in-situ analytical instruments, or collecting duplicate samples for analysis in the mobile field laboratory, to assess random errors originating from sampling and analytical procedures, and to estimate the uncertainty of measurements. Since, NORISC is dealing with the assessment of contamination of small-size areas within cities, and relies mainly on in-situ analytical methods, a cost-effective technique, using robust analysis of variance for the estimation of necessary quality control parameters, and measurement uncertainty, is explained with examples.

Measurement uncertainty in the interpretation of contaminated land may have profound effects on the realistic assessment of the extent of contamination, because it reduces 'misclassification'. This has legal, financial, and possibly health implications, from both the unnecessary remediation of “uncontaminated” land, and parts of the property regarded as uncontaminated that are in fact contaminated.

The recommended technique separates the different components of variance (sampling, analytical, geochemical), and indicates which ones are unacceptably high, and may need improvement. Further, it estimates the combined measurement uncertainty, and its use in the probabilistic classification of contaminated land into: (a) uncontaminated, (b) possibly contaminated, (c) probably contaminated, and (d) contaminated. Thus, improving the reliability of classification of contaminated land, compared to the traditional deterministic geochemical approach.

Regarding the use of geostatistics for the estimation of uncertainty in the NORISC approach, there is the question of collection of an adequate number of samples for the valid calculation of the necessary geostatistical parameters. However, the use of two times the kriging standard error of estimation as an estimate of uncertainty is explained, and its utilisation in the probabilistic classification of contaminated and uncontaminated land.

CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>Estimation of uncertainty due to sampling and analysis</td>
<td>3</td>
</tr>
<tr>
<td>Practical detection limit and analytical precision</td>
<td>6</td>
</tr>
<tr>
<td>Handling of values below detection limit</td>
<td>7</td>
</tr>
<tr>
<td>Limitations of the geochemical data set: quality &amp; reliability</td>
<td>7</td>
</tr>
<tr>
<td>Effects of uncertainty and probabilistic risk assessment maps</td>
<td>8</td>
</tr>
<tr>
<td>Example 1: Classification of contaminated land with an uncertainty of 10%</td>
<td>10</td>
</tr>
<tr>
<td>Example 2: Probabilistic risk assessment mapping</td>
<td>10</td>
</tr>
<tr>
<td>Geostatistics: two times the kriging standard error of estimation as an estimate of uncertainty</td>
<td>11</td>
</tr>
<tr>
<td>Discussion and conclusions</td>
<td>13</td>
</tr>
<tr>
<td>References</td>
<td>14</td>
</tr>
<tr>
<td>Instructions for downloading program ROBCOOP4.EXE</td>
<td>16</td>
</tr>
</tbody>
</table>
QUALITY ASSURANCE AND QUALITY CONTROL (QA/QC) FOR IN-SITU GEOCHEMICAL METHODS, ESTIMATION OF MEASUREMENT UNCERTAINTY AND CONSTRUCTION OF PROBABILITY RISK ASSESSMENT MAPS

Alecos Demetriades and Helen Karamanos
Institute of Geology and Mineral Exploration, Greece

INTRODUCTION

Estimation of both random and systematic errors in geochemical analysis and measurements for environmental studies has become an established part of good professional practice, and are, therefore, included in the routine field geochemical analyses and measurements of the NORISC approach.

The separate estimation of sampling and analytical errors using analysis of variance (ANOVA) was pioneered for geochemical surveys in mineral exploration by Miesch (1964, 1967, 1973, 1976), Garrett (1969, 1973, 1983) and Garrett and Goss (1979), and has been described for pollution investigations by Holcombe (1988). Combination of robust ANOVA and sampling and analytical quality control has been applied to the evaluation of precision requirements for applied geochemistry (Ramsey, 1992), and the estimation of measurement uncertainty (Ramsey, 1997, 1998; Ramsey and Argyraki, 1997). The robust analysis of variance technique proposed by Ramsey (1998) requires a minimum number of eight (8) locations at which to take duplicate samples or measurements, in order to give a reasonably reliable estimate of component variances (sampling, analytical & geochemical). According to Ramsey (1992), and all researchers working on contaminated sites, the application of these techniques to environmental surveys is particularly appropriate due to the high degree of heterogeneity, often associated with anthropogenic contamination of the environment.

The robust analysis of variance technique proposed by Ramsey (1998) is adapted for the purposes of the NORISC field geochemical methods, and requires two measurements to be taken for each determinand (Fig. 1):

(a) at the routine location, and
(b) at the duplicate location (to be at a distance of 1 to 2 m from the routine location).

OR in case physical samples are collected, two analyses for each determinand should be performed on either two sub-samples of the routine and duplicate samples or, at worse, each of the routine and duplicate samples to be analysed two times in a random manner (Fig. 2).
Fig. 2. Balance hierarchical geochemical sampling and analytical scheme for the estimation of geochemical, sampling and analytical variance and random components of measurement uncertainty by the in-situ geochemical methods used in the NORISC approach. This is the case when physical samples are collected and analysed by the in-situ methods. In case the routine and duplicate samples from the same location cannot be split into two sub-samples for analysis, then each routine and duplicate sample is analysed two times.

As matter of logistics, the robust analysis of variance technique is not so demanding, because it requires a minimum of 8 quadruple analyses or measurements, i.e.,

- 8 routine locations x 2 analyses or measurements, and
- 8 duplicate locations x 2 analyses or measurements,

making a total of 32 analyses or measurements for each determinand. Whereas the classical analysis of variance technique proposed by Garrett (1969, 1973, 1983) requires a minimum of 30 duplicate analyses or measurements, i.e.,

- 30 routine locations x 2 analyses or measurements, and
- 30 duplicate locations x 2 analyses or measurements,

making a total of 60 analyses or measurements for each determinand.

Since, the robust analysis of variance technique proposed by Ramsey (1998) is cost-effective compared to other techniques, it is recommended to be applied to all NORISC field geochemical surveys on contaminated land (Figs. 1 & 2).

The following notation should be followed:

(a) routine measurement at a certain location: location or sample number to be followed by A
(b) repeat measurement at the routine location: location or sample number to be followed by B
(c) duplicate measurement at the same location (1-2 m away from routine location): location number or sample number to be followed by DA
(d) repeat measurement at duplicated location: location or sample number to be followed by DB.

ESTIMATION OF UNCERTAINTY DUE TO SAMPLING AND ANALYSIS

As pointed out by Ramsey (1998) two of the component variances can be classed as measurement uncertainty, and these are the sampling \( s^2_{\text{samp}} \) and analytical variance \( s^2_{\text{anal}} \). The third component is the between location variance, due to real variation of the determinand across the investigated site. This is called the geochemical variance \( s^2_{\text{geochem}} \), in this particular case of a geochemical investigation.

**Sampling uncertainty**, or within-location variance, is partially due to small scale geochemical variation within the location, but represents the uncertainty in all samples or in-situ
measurements that can be taken or measured from that particular “location”, as specified by the investigation, i.e., one or two metre radius, depending on the distance of grid nodes.

All three variances of a particular determinand in a material, such as soil or ground water, can be summed up to give the total variance ($s_{total}^2$) of a survey. This is the figure that would be estimated when calculating the standard deviation of all analyses or measurements, and can be expressed by:

$$s_{total}^2 = s_{geoch}^2 + s_{samp}^2 + s_{anal}^2$$ \hspace{1cm} (1)

Ramsey et al. (1992) proposed the term technical variance ($s_{tech}^2$) for the sum of the sampling, $s_{samp}^2$, and analytical, $s_{anal}^2$, variance of a particular determinand in a material, such as soil or ground water, i.e.,

$$s_{tech}^2 = s_{samp}^2 + s_{anal}^2$$ \hspace{1cm} (2)

Hence, the total variance ($s_{total}^2$) of a particular determinand in a material, such as soil or ground water, becomes:

$$s_{total}^2 = s_{geoch}^2 + s_{tech}^2$$ \hspace{1cm} (3)

The measurement uncertainty ($u$) can be estimated using this “bottom up” approach, from the combination of sampling and analytical variance, giving the measurement variance as:

$$\text{measurement uncertainty, } u = (s_{meas}) = \sqrt{(s_{samp}^2 + s_{anal}^2)}$$ \hspace{1cm} (4)

It is a normal statistical procedure to increase the confidence interval of the uncertainty by multiplying by a coverage factor ($k$) 1.96 (for the 95% confidence level) to give the expanded or extended uncertainty ($U$):

$$\text{expanded uncertainty, } U = k.u = 1.96.s_{meas}$$ \hspace{1cm} (5)

The expanded or extended uncertainty ($U$) expressed as a percentage in relation to the mean concentration of a particular determinand gives the relative uncertainty ($U\%$):

$$\text{relative uncertainty, } U\% = \frac{196.s_{meas}}{m}$$ \hspace{1cm} (6)

where:

$m$ is the estimated mean concentration of a determinand at the investigated site.

The calculated value of the uncertainty applies to measurements on single samples or single locations taken or measured during the investigation. If $n$ multiple samples or in-situ measurements are taken at any individual location within the investigated site, the uncertainty on the average for that location is the value given by equation (6) divided by $\sqrt{n}$. This is equal to the standard error on the mean value ($s/\sqrt{n}$). For example, the estimated relative uncertainty at a location where four measurements ($1A$, $1B$, $2A$, $2B$) have been made would be half ($1/\sqrt{4}$) of the value as given by equation (6), which is modified as shown below:

$$\text{relative uncertainty, } U\% = \left(\frac{196.s_{meas}}{m}\right) \div 2$$ \hspace{1cm} (7)
The upper limit of expanded uncertainty, \( U \), at the 95\% confidence level is estimated by the equation:

\[
c + U = c(1 + \frac{U\%}{100})
\]

and the lower limit of expanded uncertainty is calculated by the equation:

\[
c - U = c(1 - \frac{U\%}{100})
\]

where:
- \( c \) = the concentration of the determinand in the sample medium or at the location
- \( U \) = the expanded uncertainty at the 95\% confidence level
- \( U\% \) = the relative expanded uncertainty at the 95\% confidence level

The above equations (8) and (9) may be refined if the analytical bias is estimated by the use of certified reference samples (Ramsey and Argyraki, 1997; Ramsey, 1998), which is a procedure employed by conventional laboratories. Although it may seem impractical, the analysis of certified reference materials can be carried out by the field laboratory.

According to Ramsey and Argyraki (1997) the uncertainty interval of any concentration, \( c \), becomes asymmetric. The upper limit can, therefore, be calculated from:

\[
c + U = c(1 + \frac{U\%}{100})\left(1 - \frac{B_a}{100}\right)
\]

and the lower limit from:

\[
c - U = c(1 - \frac{U\%}{100})\left(1 - \frac{B_a}{100}\right)
\]

where:
- \( c \) = the concentration of the determinand in the sample medium or at the location
- \( U \) = the expanded uncertainty at the 95\% confidence level
- \( U\% \) = the relative expanded uncertainty at the 95\% confidence level
- \( B_a \) = the analytical bias estimated as a percentage by regression

Ramsey and Argyraki (1997) point out that the interpretation of relative uncertainty in the measurements of a particular determinand in soil, for example, assumes that it does not vary with concentration. Such a case, has been observed in determinands, the analytical precision of which is considerably higher than the detection limit (Thompson and Howarth, 1976, 1978). Since the relative analytical precision \((P_{c\%})\) varies according to the concentration of the determinand the above equation (6) may be improved, by incorporating precision, estimated on survey samples (Ramsey 1997, 1998; Ramsey and Argyraki, 1997).

\[
c + U = c\left(1 + \frac{U\%}{100}\right)\left(1 - \frac{P_{c\%}}{100}\right)
\]

and the lower limit of expanded uncertainty is calculated by the equation:

\[
c - U = c\left(1 - \frac{U\%}{100}\right)\left(1 - \frac{P_{c\%}}{100}\right)
\]
The practical detection limit, and analytical precision, can easily be estimated by the field laboratory using the method described below.

PRACTICAL DETECTION LIMIT AND ANALYTICAL PRECISION

The practical detection limit and analytical precision can be estimated by using the method proposed by Howarth and Thompson (1976) and Thompson and Howarth (1976, 1978), with modifications made by the first author. Replicated analysis in the field or conventional laboratory is performed for at least 50 randomly selected samples. The steps followed are:

1. Calculate the mean values of the 50 pairs \([X_i + X_j]/2\). According to Thompson and Howarth (1978) this mean value is an estimate of true concentration of an element for the particular analytical method used.
2. Calculate the absolute differences between each pair \(|X_i - X_j|\). The absolute difference is an estimate of the standard deviation, \(\sigma_c\), at that particular concentration. \(|X_i - X_j|\) is normally distributed and relates to the parent population, with a standard deviation \(\sigma_c\), such that: \(\sigma_d = \sqrt{2}\sigma_c\), where \(\sigma_d\) is the standard deviation of the difference \([X_i - X_j]\);
   \[d = 1.128\sigma_c\], where \(d\) is the mean value for the difference; and
   \[M_d = 0.954\sigma_c\], where \(M_d\) is the median value for the difference. \(\sigma_c\) can be obtained from each of these relationships, but the median \((M_d)\) is the most convenient estimator, because it is (i) relatively little affected by wild values; (ii) readily estimated graphically, and (iii) corresponds very closely to \(\sigma_c\) without further calculation (Fletcher, 1981).
3. Arrange list in increasing order of concentration means;
4. From the first 11 results calculate the mean concentration (Group mean) and the median difference (Group median);
5. Repeat step 4 for each successive group of 11 samples, ignoring any remainder less than 11;
6. Calculate the linear regression of the median difference (y-axis, dependent variable) on the means (x-axis, independent variable). At this point a modification has been introduced by the first author. In classical regression, \((Y = a + bX)\), a linear relationship is quantified by fulfilling the following requirements of (a) dependency and (b) knowing one variable without error. Thompson and Howarth (1978) assumed that the group means are the independent variable or predictor (X), by which the group median difference (Y) is estimated. The question posed is the following: which is really the dependent variable? Since, both variables are derived from the grouping of the same analytical data set, they are subject to errors of the same order of magnitude. It is concluded, therefore, that the requirements of classical regression cannot be met. To overcome this situation Kermack and Haldane (1950) developed the reduced major axis line, which is the line of best-fit between a set of points (Till, 1974). Essentially, is the best-fit line between the two regression lines of \((Y = a + bX)\) and \((X = a + bY)\). Hence, errors of estimation are minimised;
7. Obtain from the major axis regression line of the group median differences, \([X_1 - X_2]\), on the group means, \([X_i + X_j]/2\], the intercept, \(a\), and coefficient, \(b\).
8. Multiply by 1.048 (i.e., 1/0.954) the intercept, \(a\), and coefficient, \(b\), to obtain \(\sigma_o\) and \(k\) respectively; from the regression \(\sigma_c = \sigma_o + kc\), so that the precision, \(P_c\), is given by
   \[P_c = [(1.96\sigma_o/X_o) + (1.96k)]\] (14)
   which is the variation at the two standard deviation (95%) confidence level.
9. Calculate the percentage precision \((P_c\%)\) by using the equation:
   \[P_c\% = [(1.96\sigma_o/X_o) + (1.96k)] .100 = [(196\sigma_o/X_o) + (196k)]\] (15)
where \( X_i \) is the element concentration determined on individual samples. Hence, it is possible to estimate, by this method, the precision for every determination.

10. Calculate the detection limit. Detection limit is normally defined as the concentration that gives rise to a signal equal to twice the standard deviation of blank fluctuations, i.e., at a value of \( P_c = 100\% \) and \( X_d = 1.96\sigma_o \). At concentrations higher than the detection limit, precision falls asymptotically towards the value of \( 1.96k \) as defined in the expression \( P_c = [(1.96 \cdot \sigma_o/X_d) + (1.96k)] \). For further information, and the implications involved in the estimation of these quality control parameters, Thompson and Howarth (1976) should be consulted.

It is important to understand the asymptotic nature of precision, and that it is wrong to quote a single value for precision, i.e., at concentrations higher than the detection limit, precision falls asymptotically towards the value of \( 1.96k \) or \( 196k \) in the above expressions (refer to Fletcher, 1981, Fig. 2-5, p.32). On the geochemical distribution maps the relative precision equation should be given, so the reader can estimate precision at any specific concentration.

Practical detection limits determined by this method are subject to the variation of element concentrations in the selected random samples. In case the samples have a good distribution of element concentrations, approaching a normal Gaussian distribution, the practical detection limits of these elements are either the same or very close to instrument detection limits. Elements which have non-Gaussian distributions, their practical detection limits are normally very different from those quoted by the analysts.

Finally, the samples or locations selected for replicate measurements should include very low, low, moderate, high and very high concentrations of the determinands studied. Hence, this selection can only be made upon completion of the routine site investigation.

**HANDLING OF VALUES BELOW DETECTION LIMIT**

For statistical purposes, it is customary in geochemical surveys, to set all concentration values quoted as “below detection limit” to half the reported limit.

**LIMITATIONS OF THE GEOCHEMICAL DATA SET: QUALITY & RELIABILITY**

The combined sampling and analytical variance (the technical variance), according to Ramsey et al. (1992) and Ramsey (1993, 1998), should not exceed the upper limit of 20% of the total variance. An upper limit for analytical variance is set at 4% of the total variance. Hence, the sampling variance should not exceed the upper limit of 16% of the total variance.

It is significant, as pointed out by all researchers in this field, for the greatest part of the variance to be ascribed to the geochemical data variance or geochemical (spatial) variation, otherwise a contoured geochemical distribution map cannot be produced (Garrett, 1969; Howarth, 1983; Sharp, 1987). In the latter case, where geochemical data show no spatial persistence, element concentrations may be plotted at the sample sites as variable-size dots (Bølviken et al., 1986; Björklund and Gustavsson, 1987; Lahermo et al., 1990), symbols (De Vos et al., 1996), Exploratory Data Analysis symbols (Englund and Sparks, 1988; O’Connor et al., 1988; Demetriades, 1990), or recording on maps the sample site analytical data (Van der Sluys et al., 1997). However, final decisions about geochemical distribution map plotting may be decided upon, following a thorough geostatistical structural analysis of the data (Davis, 1973; Miesch, 1975; Journel and Huijbregts, 1978; Rendu, 1978; Clark, 1979; Howarth, 1983; Sharp, 1987; Isaaks and Srivastava, 1989; Reimann, 1998). According to Sharp (1987, p.11) to construct a valid geochemical contour map from point data, two very specific rules should be satisfied:

- **Rule 1:** The data must show spatial persistence (autocorrelation) up to the second nearest neighbours. The rule insists that the data show sufficient continuity that a minimum determination of both the slope, and its general curvature, can be obtained for any interpolation.
- **Rule 2**: The contour interval shall be selected so that the probable error of any point does not exceed one-half of a contour interval. This rule insists that any new point randomly selected between the drawn contour lines has at least an even chance of being valid.

Ramsey *et al.* (1992) and Ramsey (1993) stressed that, application of ANOVA and Robust ANOVA techniques to environmental surveys, is particularly appropriate due to the high degree of heterogeneity often associated with anthropogenic contamination of the environment. They suggested the graphical display of data quality parameters in the form of a pie chart (Fig. 3). Visual representation of variance on all element distribution maps, gives the reader direct access to significant information about the quality and reliability of geochemical data.

![Pie chart showing proportions of variance contributed by measurement processes of sampling and analysis for e.g. Pb in soil at an investigated site.](image)

Fig. 3. Example of the proportions of variance contributed by measurement processes of sampling and analysis for e.g. Pb in soil at an investigated site. For geochemical information to be reliable, the maximum analytical and sampling variance should be <4% and <16% respectively. The pie chart shows total Pb data to be within the specifications, and the measurement variance and, hence, uncertainty is dominated by the contribution from the sampling (18.73%), rather than from the chemical analysis (0.09%) (Demetriades, 1999a, Fig. 2C.1, p. 45).

Apart from the estimation of sampling, analytical and geochemical variability, it is possible to estimate the practical detection limit and analytical precision for each element or determinand at each site, as has already been described (Howarth and Thompson, 1976; Thompson and Howarth, 1976, 1978; Fletcher, 1981; Demetriades, 1999a). However, for the purposes of some sites this may not possible, due to the minimum number of random duplicate measurements that must be made, *i.e.*, a minimum of at least 50 duplicate measurements. In the case of large sites, it is advisable to use this procedure.

### EFFECTS OF UNCERTAINTY AND PROBABILISTIC RISK ASSESSMENT MAPS

In order to address the uncertainty of the measurements, a probabilistic classification of contaminated land (Ramsey and Argyraki, 1997) is used to produce *probabilistic hazard or risk assessment maps*. The probabilistic hazard or risk assessment maps address decision rule uncertainty by considering the uncertainty as being on the statutory threshold or guideline value used to determine the decision rule. The classification defines four categories based on the extent of overlap of uncertainty with a single threshold or guideline value (Table 1):

1. **Uncontaminated**,  
2. **Possibly contaminated**,  
3. **Probably contaminated**, and  
4. **Contaminated**.

For the category “uncontaminated”, for example, the entire range of uncertainty is lower than the regulatory threshold or guideline value (*T*). The probability of this site being contaminated, because the element concentration lies over the statutory threshold is, therefore, <0.025 (*i.e.*, 2.5%). This computational device can be used to classify soil or ground water samples directly, without the need to calculate uncertainty values for each measurement.
The use of duplicate samples, as has been shown above, is one method that has been proposed for the estimation of sampling uncertainty (Ramsey and Argyraki, 1997).

Table 1. Probabilistic classification of contaminated soil with measured concentration \(C\) and uncertainty \(U\), based on the probability that the contaminant concentration is greater than a particular threshold or guideline value \(T\) (From Ramsey and Argyraki, 1997, Table 3, p. 251).

<table>
<thead>
<tr>
<th>Classification</th>
<th>Concentration range with uncertainty on (C)</th>
<th>Concentration range with uncertainty on (T)</th>
<th>Probability of (C &gt; T), i.e., (C) being over threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncontaminated</td>
<td>(C + U &lt; T)</td>
<td>(C &lt; T - U)</td>
<td>&lt;0.025</td>
</tr>
<tr>
<td>Possibly contaminated</td>
<td>(C &lt; T &lt; C + U)</td>
<td>(T - U &lt; C &lt; T)</td>
<td>0.025-0.500</td>
</tr>
<tr>
<td>Probably contaminated</td>
<td>(C - U &lt; T &lt; C)</td>
<td>(T &lt; C &lt; T + U)</td>
<td>0.500-0.975</td>
</tr>
<tr>
<td>Contaminated</td>
<td>(T &lt; C - U)</td>
<td>(C &gt; T + U)</td>
<td>&gt;0.975</td>
</tr>
</tbody>
</table>

\(\text{NOTE: The effect of uncertainty on the concentration estimate is evaluated (column 2), but a rapid intermediate calculation can consider the uncertainty as being on the threshold or guideline value (column 3). Uncertainty values are stated for 95% confidence interval, but could be recalcualted for whatever confidence interval is considered appropriate.}\)

Ramsey and Argyraki (1997) also mention that there is evidence of uncertainty changes with the concentration. The *upper limit of uncertainty* for each single concentration measurement is given by:

\[
C + U = C \left(1 + \frac{U\%}{100}\right) \quad (16)
\]

where:
- \(C\) is the concentration of a determinand in soil or ground water,
- \(U\) is the expanded measurement uncertainty, and
- \(U\%\) is the uncertainty relative to the mean concentration.

The *lower limit of uncertainty* is similarly given by:

\[
C - U = C \left(1 - \frac{U\%}{100}\right) \quad (17)
\]

A “short cut” can be used by classifying geochemical samples directly, without calculating uncertainties for each measurement, i.e., the uncertainty values for the particular measurements made are calculated for the concentration equal to the statutory threshold or guideline value \(T\). When the “short cut” method is used, with the uncertainty expressed on the threshold or guideline value, then the upper and lower limits are given by:

\[
\text{Upper limit: } \quad T + U = T \left(1 + \frac{U\%}{100}\right) \quad (18)
\]
\[
\text{Lower limit: } \quad T - U = T \left(1 - \frac{U\%}{100}\right) \quad (19)
\]

The probabilistic classification boundaries can then be calculated using these equations:

\[
\text{Uncontaminated: } \quad C = T - U \text{ or } C + U = T \quad (20)
\]
\[
\text{Contaminated: } \quad C = T + U \text{ or } C - U = T \quad (21)
\]

**Example 1: Classification of contaminated land with an uncertainty of 10%**

The equations used for the classification of contaminated land using 10% measurement uncertainty on the ICRCL (1987) trigger concentration of 500 mg Pb/kg soil for domestic gardens
and allotments, and the classification of land into: (i) uncontaminated, (ii) contaminated, (iii) possibly contaminated and (iv) probably contaminated are given below:

Uncontaminated:  
\[ C + U = T \]  
\[ C + 0.1 C = 500 \]  
\[ C = \frac{500}{1.1} = 454.5 \text{ mg Pb/kg soil} \]  \[(22)\]

Contaminated:  
\[ C - U = T \]  
\[ C - 0.1 C = 500 \]  
\[ 0.9 C = 500 \]  
\[ C = \frac{500}{0.9} = 555.6 \text{ mg Pb/kg soil} \]  \[(23)\]

Possibly contaminated:  
\[ 454.5 < C < 500 \text{ mg Pb/kg soil} \]  \[(24)\]

Probably contaminated:  
\[ 500 < C < 555.6 \text{ mg Pb/kg soil} \]  \[(25)\]

**Example 2: Probabilistic risk assessment mapping**

The example below (Table 2, Figs. 4 & 5) shows the procedure used to estimate the analytical, sampling and geochemical variance, as well as the measurement uncertainty at the 95% confidence level, employing the technique proposed by Ramsey (1998). The public domain program ROBCOOP4 is used for the estimation of the component variances by robust statistics (Table 2). The Ramsey (1998) paper: “Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources”, and the program ROBCOOP4 can be downloaded from the Web site of the Journal of Analytical Atomic Spectrometry [http://www.rsc.org/jaas].

Table 2. Repeat analytical results of mercury (ppb) on routine and duplicate sub-samples, robust statistics, and values of uncertainty and expanded uncertainty (%) at the 95% confidence level according to the method proposed by Ramsey (1998).

<table>
<thead>
<tr>
<th>Routine sub-samples</th>
<th>Duplicate sub-samples</th>
<th>Hg_1A</th>
<th>Hg_1B</th>
<th>Hg_2A</th>
<th>Hg_2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>4A</td>
<td>4B</td>
<td>40</td>
<td>50</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>20A</td>
<td>20B</td>
<td>80</td>
<td>80</td>
<td>60</td>
<td>70</td>
</tr>
<tr>
<td>26A</td>
<td>26B</td>
<td>810</td>
<td>790</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>37A</td>
<td>37B</td>
<td>2330</td>
<td>2370</td>
<td>2590</td>
<td>2690</td>
</tr>
<tr>
<td>44A</td>
<td>44B</td>
<td>8970</td>
<td>10300</td>
<td>1695</td>
<td>1550</td>
</tr>
<tr>
<td>51A</td>
<td>51B</td>
<td>600</td>
<td>640</td>
<td>1590</td>
<td>1420</td>
</tr>
<tr>
<td>60A</td>
<td>60B</td>
<td>220</td>
<td>240</td>
<td>430</td>
<td>570</td>
</tr>
<tr>
<td>82A</td>
<td>82B</td>
<td>20</td>
<td>25</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>98A</td>
<td>98B</td>
<td>30</td>
<td>25</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>118A</td>
<td>118B</td>
<td>20</td>
<td>25</td>
<td>20</td>
<td>25</td>
</tr>
</tbody>
</table>

**Robust analysis of variance results**

<table>
<thead>
<tr>
<th>Statistical parameters</th>
<th>Geochemical</th>
<th>Sampling</th>
<th>Analytical</th>
<th>Total</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>442004.25</td>
<td>75842.96</td>
<td>317.93</td>
<td>518365.13</td>
<td></td>
</tr>
<tr>
<td>Standard deviation (+/-)</td>
<td>664.83</td>
<td>275.40</td>
<td>22.76</td>
<td>719.98</td>
<td>500.40</td>
</tr>
<tr>
<td>Variance (%)</td>
<td>85.27</td>
<td>14.63</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Uncertainty 138.17

Expanded uncertainty (%) 54.12 at the 95% confidence interval
An alternative approach is the construction of probabilistic hazard or risk assessment maps by using the kriged estimates of the studied determinant. Kriging is a geostatistical technique using the kriged estimates of the studied determinant. Kriging is a geostatistical technique.

GEOSTATISTICS: TWO TIMES THE KRIGING STANDARD ERROR OF ESTIMATION AS AN ESTIMATE OF UNCERTAINTY

An alternative approach is the construction of probabilistic hazard or risk assessment maps by using the kriged estimates of the studied determinant. Kriging is a geostatistical technique.
used to estimate determinand concentrations at and between sampling locations (Davis, 1973; Miesch, 1975; Journel and Huijbregts, 1978; Rendu, 1978; Clark, 1979; Krige, 1984; Isaaks and Srivastava, 1989; Demetriades, 1999b). It is based on the characterisation of the change in covariance of measurements made at pairs of sampling locations, with increasing distances of separation. The model fitted to this variation is called the variogram. The intercept of the variogram on the y-axis, called the ‘nugget’, is the variance remaining at zero distance (x-axis), and reflects random components of the measurement uncertainty.

The construction of a satisfactory variogram depends on the collection of a large number of samples, spread evenly across the property. This is indeed a problem when small areas in the order of one hectare are studied. The method proposed by Ramsey (1998) is much more cost-effective and adaptable to small surveys.

Nevertheless, if an adequate number of samples is collected, and a valid geostatistical structural analysis is performed using variograms, and the extracted geostatistical parameters are cross-validated by point kriging, i.e., a process by which the actual point values are estimated from the surrounding ones using the geostatistical parameters obtained from the variogram. The objective is to minimised the difference between estimated kriged and actual values.

Kriging can be used to estimate not only the concentration of intermediate locations within the sampling area, but also a kriging standard error of estimation, which is also an uncertainty (Ramsey, 1998; Tristán et al., 1999). The kriging standard error of estimation can be used, therefore, to estimate the uncertainty of each kriged estimate generated by point or block kriging (Demetriades, 1999b; Tristán et al., 1999). The essential difference between the geostatistical uncertainty, and the robust statistical method proposed by Ramsey (1998) is that it refers to concentrations between rather than at the sampling locations. A comment made by Ramsey (1998) is that if the method he proposes is used to improve the estimate of uncertainty at the sampling locations (e.g., by including sampling bias), then kriging can be used to refine the estimates of uncertainties between the sampling locations. In this way the two approaches are complementary in their objectives.

Since, geochemical data are normally positively skewed (Ahrens, 1954a, 1954b), it is recommended to study the statistical distribution of each determinand to begin with, and transform the results, if necessary, to logarithms base-10 (log_{10}) or base-e (log_{e}, natural logarithms) before using geostatistics.

Two times the kriging standard error of estimation (ksd) is used as an estimate of uncertainty, “U”. It is noted that in point and block kriging, the kriging standard error of estimation is different for each point or block of the study area. Therefore, the uncertainty is estimated for each point or block. As explained earlier, the uncertainty is expressed on the threshold or guideline value, as a shortcut to define “contaminated” and “uncontaminated” land.

The equations using two times the kriging standard error of estimation (ksd) as a measure of uncertainty on the threshold or guideline (T) concentration of a specific determinand, and the classification of land into “uncontaminated” and “contaminated” are given below:

\[
\text{Uncontaminated:} \quad \text{kriged } C + (2 \times \text{ksd}) = T \quad \text{(26)}
\]

\[
\text{or}
\]

\[
\text{Contaminated:} \quad \text{kriged } C - (2 \times \text{ksd}) = T \quad \text{(27)}
\]

\[
\text{or}
\]

\[
\text{kriged } C = T - (2 \times \text{ksd})
\]
Using the above relations the classification of contaminated land into “possibly” and “probably contaminated” can be made:

### Possibly contaminated:
\[
[T - (2 \times ksd)] < \text{kriged } C < T
\]

### Probably contaminated:
\[
T < \text{kriged } C < [T + (2 \times ksd)]
\]

**DISCUSSION AND CONCLUSIONS**

The robust statistical method proposed by Ramsey (1998) for the estimation of sampling, analytical and geochemical variance, as well as measurement uncertainty is cost-effective, especially for the investigation of contaminated land in the order of one hectare. It is a simple method and requires the collection of duplicate samples from a minimum of eight locations, randomly distributed over the investigated site. Geostatistics can be used, but to obtain an effective variogram model a large number of samples is required. Therefore, geostatistics can only be used in the investigation of large areas, and as mentioned above as a complementary method. Disjunctive kriging (Rivoirand, 1994) is another geostatistical technique that estimates the risk that the real value at a site exceeds a prescribed threshold.

Once the uncertainty of measurements has been estimated for a particular survey, the question that arises is whether the particular level of uncertainty is acceptable. In applied geochemical surveys the analytical quality control in trace element concentrations is set, for example, at 10% analytical precision, as common practice. Such targets are rarely based, however, on a rigorous evaluation of the precision required for the interpretation of the analytical results. When the requirement is for the interpretation of concentration measurements Ramsey et al. (1992) suggested that the limit for measurement variance and, therefore, uncertainty is given by:

\[
\frac{s^2_{\text{meas}}}{s^2_{\text{total}}} < 20\%
\]

When this proportion is exceeded, then the identification of “hot spots” on contaminated land tends to become progressively less reliable. Such a level is related to the variability of the environment under study (e.g., the degree of contamination of the property), rather than using a level fixed in relative terms (e.g., 10% RSD). No matter how large the measurement uncertainty becomes, however, it is always possible to use the estimated uncertainty values. These can be used to test whether one location is contaminated more than another location (or more than a threshold or guideline value), using traditional statistical techniques. In that sense, therefore, there is no universal limit to the acceptable level of uncertainty. Nevertheless, it may prove that the uncertainty of a certain survey prevents the quantification of a particular objective, such as the delineation of a contamination “hot spot”. In such a case, a second survey would be required, which would have the value of uncertainty reduced to an appropriate level in order to define contamination. This acceptable level of uncertainty needs to be calculated from the results of the first survey.

As pointed out by Ramsey and Argyraki (1997) the use of the probabilistic classification of contaminated land can have large financial implications for its treatment. The specification, for example, of 95% confidence for the classification of the land as “contaminated” may mean that the cost of remediation is far higher than if the traditional deterministic classification is used (actual values). However, the use of the deterministic method would lead to a high probability of some contaminated areas being misclassified as “uncontaminated” and, therefore, will not be considered for treatment. Consequently, the traditional approach using the raw values of contaminants has even greater financial implications than the probabilistic, in the longer term with respect to legal liability.
According to Ramsey and Argyraki (1997, p.255) the advantages of the probabilistic approach are:

1. A reduced risk of misclassification of land (e.g., cost savings on either unnecessary remediation, or overlooking the need for remediation).
2. Variable probabilities of misclassification can be incorporated, to reflect the particular circumstances of the site (e.g., intended land use).
3. Cost-benefit analysis can be employed on the design of the sampling survey. For example, in assessing the effectiveness of taking composite samples for reducing uncertainty.
4. Uncertainty estimates can also be incorporated in environmental risk assessment to give uncertainties in estimated risks, and
5. Facilitates identification of causes of measurement uncertainty, and indicates where improvements need to be made (e.g., in sampling rather than in the analytical method).

Disadvantages of the probabilistic classification of contaminated land are:

(a) More land will become classified as “possibly contaminated” than is considered “contaminated” under the deterministic method. Such areas can, however, be re-sampled by a method with lower uncertainty, to clarify its true status.

(b) Increased cost of initial survey, typically +10% to +20% for duplicate samples. The measurement of uncertainty has certainly a cost, but the information gained will be more useful than unqualified measurements of concentration.

(c) Increased complexity in explaining the survey findings. However, even the general public is becoming more familiar with measurements expressed in a probabilistic way, for example, the weather forecasts, and

(d) The estimates of uncertainty made for sampling methods are site specific, and are not generally applicable, as would be the case for uncertainty from analytical methods. Given that the classification of contaminated land is moving towards a more site-specific approach this may not be a serious limitation.

Quality assurance systems are nowadays an absolute necessity to ensure that organisations engaged in field sampling provide unbiased estimates of measurement uncertainty, as well as for the concentrations of contaminants. Therefore, the proposed scheme should be used in the NORISC approach in the production of scientifically reliable and legally defensible results.

REFERENCES

Demetriades, A., 1999b. Data processing: statistical and geostatistical methods and geochemical map production. In: A. Demetriades (editor), Geochemical atlas of the Lavrion urban area for environmental protec-


INSTRUCTIONS FOR DOWNLOADING PROGRAM ROBCOOP4.EXE:

1. Link to the internet site of the Journal of Analytical Atomic Spectrometry at http://www.rsc.org/
2. Select “A-Z Contents”
3. Select “Electronic Supplementary Information”
6. Download (a) ROBCOOP4.EXE program, (b) Sample data and (c) Full Article: “Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources” by M.H. Ramsey.

If problems are encountered during downloading, contact Alecos Demetriades at either igmegeox@otenet.gr OR alecos_demetriades@hotmail.com OR send a letter to:

Alecos Demetriades,
Institute of Geology and Mineral Exploration,
Division of Geochemistry & Environment,
70 Messoghion Avenue,
GR-115 27 Athens,
GREECE.

-----o-----O-----o-----o-----