CL:AIRE
technical bulletin
TB7
(March 2004)

CL:AIRE technical bulletins describe specific techniques, practices and methodologies currently being employed on sites in the UK within the scope of CL:AIRE technology demonstration and research projects. This bulletin introduces some basic statistical methods which can be used to help quantify uncertainties when managing contaminated land.

Definitions of words written in bold type may be found in the Glossary of Terms within the Publications section of the CL:AIRE Web site at http://www.claire.co.uk

Improving the Reliability of Contaminated Land Assessment using Statistical Methods: Part 1 - Basic Principles and Concepts

1. INTRODUCTION

Proper management of contaminated land requires an understanding of site conditions and of the nature and distribution of contaminants. This understanding is derived from site characterisation activities, which typically include sampling of surface and sub-surface soils and water. It is often assumed that the results obtained from these samples are representative of the actual ground and contaminant conditions but, this is not necessarily the case, due to heterogeneities in the site and uncertainties in the measurements.

This bulletin is designed to introduce some basic statistical methods which can be used to help quantify uncertainties, allow for heterogeneity, and provide confidence in making decisions when managing contaminated land.

There are two ways of approaching contaminated land assessment. One approach avoids statistical methods to make life apparently 'simple'. The other approach uses statistics to allow more reliable conclusions to be made, by taking into account the imperfections of the real world.

The first approach is initially very appealing. It is possible to apply a standard sampling protocol (e.g. BS10175, 2001) to a site and to assume that all samples taken are representative of the site. The samples can be sent to a laboratory for chemical analysis, and then an assumption made that all of the measurements of contaminant concentration in the soil are 'true'. A simple comparison between the measurements and the regulatory limit is then made to decide whether the site is 'contaminated', or needs further investigation. This approach is, however, potentially very unreliable, and can lead to serious decision errors.

Both of the assumptions made in this first approach (‘representative’ samples, and 'true' chemical analysis) are usually unjustified in most contaminated land investigations. It is possible, however, to recognise these limitations and make reliable decisions, using a few basic statistical methods. The methods can not only improve the reliability of the decisions, but can also often reduce the cost of the investigation and site development. They can be used to address the design of both the sampling and measurement strategies and the rigorous interpretation of the measurements. Statistical methods can allow for the fact that most samples are not truly representative. They can allow for variability of contamination caused by large-scale heterogeneity of contamination within sites. They also provide a rigorous tool to demonstrate that measurements of contamination are never 'true'. This inevitable uncertainty in measurements arises in three main ways. First, the small-scale heterogeneity of contamination within a sampling location causes random sampling error. Secondly, the practical interpretation and implementation of a sampling protocol can never be perfect and will tend to generate further errors. Finally, measurements in the analytical laboratory will inevitably cause additional random and systematic errors. Once these errors have been measured, the probability of making a wrong decision can be calculated, and reduced to an acceptably low level.

The objectives of this bulletin are therefore:

1. to show that the reliability of decisions made on contaminated land can be improved by application of basic statistical techniques
2. to explain how a number of basic statistical techniques can be applied - and to discuss their strengths and their weaknesses
3. to provide the basis for a second bulletin that will describe more advanced methods for contaminated land assessment, using statistics.

Fig 1. View of the potentially contaminated site in West London used as the example in this bulletin.

2. DESCRIBING CONTAMINATED LAND WITH BASIC STATISTICAL TOOLS

The best way to explain the benefits of statistics is show them being applied to a case study. An investigation of lead (Pb) contamination in top soil at a site in West London (Fig 1) demonstrates many of the common characteristics of site assessment.

The objective of this investigation was to make an initial assessment of the suitability of the site for housing development. A regular sampling grid with a 30 m spacing generated 100 locations, at which samples of top soil (0-150 mm) were taken. The concentration of total Pb, measured by a laboratory after a strong acid digest, are shown on the coordinates of the original sampling grid (Table 1).

Table 1. Measured Pb concentrations at each point on the sampling grid (mg/kg), shown by the actual coordinates used in the regular sampling grid (spacing 30 m). The degree of between-location variability is high (roughly a factor of ten). The variability within ten of these locations selected at random (i.e. A4, B7, C1, D9, E8, F7, G7, H5, I9 and J5) was also studied (Table 2). This within-location variation is substantial (e.g. a factor of two), but less than the between-location variability.

Row	A	B	C	D	E	F	G	H	I	J
1	474	287	250	338	212	458	713	125	77	168
2	378	3590	260	152	197	711	165	69	206	126
3	327	197	240	159	327	264	105	137	131	102
4	787	207	197	87	254	1840	78	102	71	107
5	395	165	188	344	314	302	284	89	87	83
6	453	371	155	462	258	245	237	173	152	83
7	72	470	194	83	162	441	199	326	290	164
8	71	101	108	521	218	327	540	132	258	246
9	72	188	104	463	482	228	135	285	181	146
10	89	366	495	779	60	206	56	135	137	149

Source: (Argyraki, 1997)

If you have any questions about this Technical Bulletin or would like further information about other CL:AIRE publications please contact us at:
CL:AIRE, 7th Floor, 1 Great Cumberland Place, London W1H 7AL Tel: 020 7258 5321 Fax: 020 7258 5322 Email: enquiries@claire.co.uk

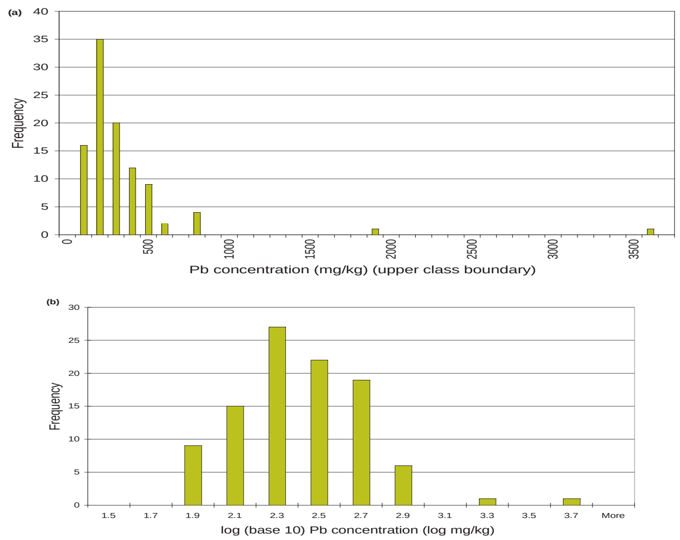

Fig 2. The Pb concentrations from the site investigation (Table 1) as a histogram of either (a) the raw concentration values, or (b) their logarithms. The frequency is the number of locations with concentration values in that particular range. The simple histogram shows a positively skewed distribution with most values being a few hundred mg/kg, but a few being thousands of mg/kg. The histogram of the 'log-transformed' values is much closer to being normally distributed, but still shows some positive skew.

The interpretation of these measurements can be helped by some basic statistical and graphical tools. How the concentration of Pb varies is much easier to see with a histogram (Fig 2). This shows the wide range of concentration, from several values under 100 mg/kg to one at 3590 mg/kg. It also shows that the distribution is not the symmetrical bell-shaped curve characteristic of a Normal frequency distribution (Fig 3), but rather is skewed towards the higher concentration values. The best estimates of the overall level of contamination can be expressed by one of three statistics; the mean, the mode or the median. The mean is simply the average (in this case 291.9 mg/kg) and is an estimate of the actual mean concentration (μ) of the whole site, which is never known. The mean concentration that is measured, $\bar{x}$, is more likely to approach the actual mean, μ, as the number of samples taken becomes larger, if there is no systematic error. The mode is roughly the most popular value (327 mg/kg) and the median is the middle value when all of the measurements are placed in ascending order (in this case 198 mg/kg). The fact that these three values differ so greatly is another sign that we are not dealing with a Normal distribution. If there is a need to convert the distribution into one that is closer to a normal distribution, then one option is to take logarithms of each concentration value (Fig 2b). The mean of this log-transformed distribution is called the geometric mean (i.e. 2.318, or 208 mg/kg), a further measure of 'central tendency'.

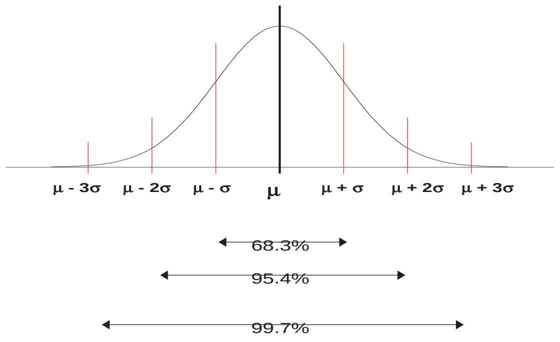

Fig 3. The Normal distribution, often used as a model to describe contaminant concentration. Ideally giving the mean (μ, estimated as $\bar{x}$) ± one standard deviation (σ, estimated as s) to describe 68.3 % of the values of contaminant concentration. Similarly $\mu \pm 2\sigma$ (estimated as $\bar{x} \pm 2s$) can be used to describe 95.4 % of the values. This is often called the 95 % confidence interval.

Three ways of expressing the scatter of the concentration values are the range, percentiles and the standard deviation. The range (e.g. 3534 mg/kg) is simply the difference between the minimum (56 mg/kg) and the maximum (3590 mg/kg) values. It is not however, a 'robust' estimate, as it is very sensitive to changes in just one value. More robust are the percentiles such as quartile ranges (e.g. bottom 25 % of values below 128.5 mg/kg, top 25 % above 327 mg/kg) or decile ranges (e.g. bottom 10 % below 82.8 mg/kg, top 10 % above 478 mg/kg). The median, quartiles and deciles are called non-parametric statistics because they do not rely on any prior assumptions about the frequency distribution. The most popular measure of scatter is the standard deviation ($s = 403$ mg/kg in this case), which is roughly the average difference between the individual measured values and the mean value. Again this is only an estimate of the standard deviation of the whole population (σ), which is never known. This does rely on the assumption that the distribution of Pb concentrations approximates to a Normal distribution (Fig 3). However, when this assumption is valid, the mean ± one standard deviation ($\mu \pm 1\sigma$) will express 68.3 % of the scatter, and $\mu \pm 2\sigma$ will give 95.4 % (Fig 3), often rounded to 95 %. Visual inspection of the experimental data (Fig 2a) shows that this particular distribution is not Normal however, so we cannot reliably apply these statistics ($\bar{x}$ and s) to this data set. Another alternative is to use robust estimators of both the mean and the standard deviation. These statistics 'accommodate' a proportion of outlying values (e.g. 10 %) that do not conform to a normal distribution. They use an iterative technique to make these estimates for the underlying distribution, and are useful for many environmental datasets (AMC, 2001).

3. CONTAMINANT CONCENTRATION AT PARTICULAR LOCATIONS

Probably more useful than describing the whole site, is to provide an estimate of the contamination concentration at individual locations within the site. This allows decisions to be made as to which parts of the site are contaminated, if any.

In the case study, we have 100 estimates of the Pb concentration at individual locations within the site (Table 1). In reality none of these estimates are the 'true' concentration of contaminant at the location. This is because samples are never perfectly representative and chemical analyses are always wrong to some extent. We can allow for this however, by quoting each of these concentration values with an estimate of its **uncertainty**. The uncertainty is defined informally as 'the range of concentration within which the true value lies with a known level of confidence'. So although we never know the true values of contaminant concentration, we know the range in which it lies, and so we can make reliable decisions if we know the uncertainty.

To estimate the uncertainty we need to know the size of the errors caused by the sampling and the analytical methods. One sort of error is **random**. It is expressed as the **precision** of the method, and it just expresses how repeatable the measurements are, with no reference to the true value. Fortunately, this error usually follows a Normal distribution, so precision can be expressed using standard deviation (s) either in units of concentration, or as precision relative to the mean value as a percentage, at 95 % confidence ($100 \times 2s/\bar{x} = 200s/\bar{x}$). The second sort of error is **systematic**. It is expressed as the **bias** of the method, and is the difference between the mean of a number of measurements by a method and the 'true' value of the concentration. The bias of an analytical method is usually estimated using a certified reference material (CRM) which has a known concentration of the substance being measured. As the true value of contaminant concentration is never known, a 'certified' value (c_{cert}) is usually used to represent it. The bias ($\bar{x} - c_{cert}$) is expressed in the units of concentration, or again relative to the mean as a percentage ($[\bar{x} - c_{cert}].100 / c_{cert}$). The term previously used to measure systematic error was accuracy. This term now has a slightly different meaning, and is applied only to a single measurement, not to a method. The **accuracy** of a single measurement is the difference between the measured value of concentration and the 'true' value, and is a quantification of the **error** of that single measurement. This difference may be caused by either random or systematic error. The difference between the meaning of these terms can be shown more clearly by using a visual analogy (Fig 4).

Precision of sampling is estimated by the taking of duplicate samples at a small proportion of the locations across the site (e.g. 10 %). These duplicate samples are not taken at exactly the same point, but at two points that are equally likely to represent the nominal location, using whatever sampling protocol and surveying equipment being applied. Therefore, if the location has a large amount of small-scale heterogeneity, the sampling precision will be high. The analytical precision is estimated by taking duplicate weighings for analysis from both of the sample duplicates, in a balanced design (Fig 5).

The analysis of these four materials for the duplicate sites (Table 2) shows that although the analytical duplicates agree quite well (e.g. <10 %), the sample duplicates often differ by up to a factor of two (e.g. location H5), showing the high level of small-scale heterogeneity at this site.

The values of the sampling and analytical precision are estimated using a technique called robust analysis of variance, described elsewhere (Ramsey, 1998). In the case study, the precision estimated for the sampling method (83.3 %, of the concentration value) is much larger than that for the analytical method (7.5 %), as suggested by the visual inspection of Table 2. This example shows a common observation, enabled

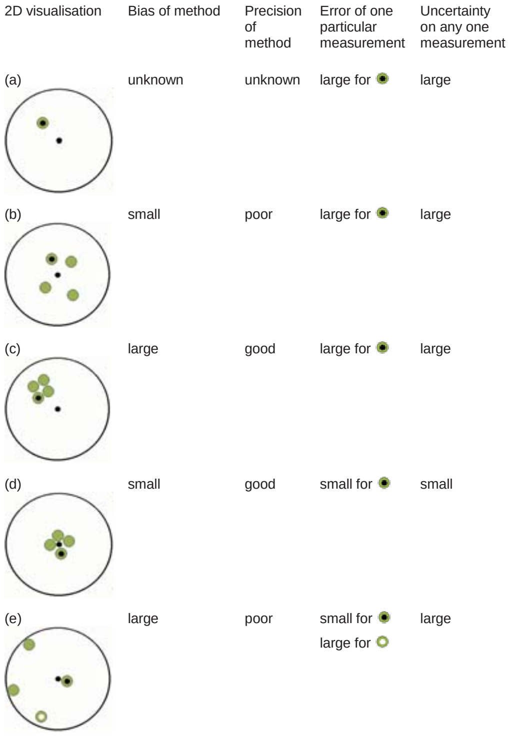

Fig. 4 Diagrammatic explanation of the difference between the terms *bias*, *precision*, *error* and *uncertainty*. (a) From a single measurement (●) of analyte concentration it is impossible to estimate the bias or precision of a method of measurement. However, the error (and accuracy) of that single measurement is simply its difference from the true value of the analyte concentration (•). (e) If the method of measurement has a large value of bias and precision, it is still possible to have a single measurement with a small error (●), but it will have large value of uncertainty. See text for further discussion.

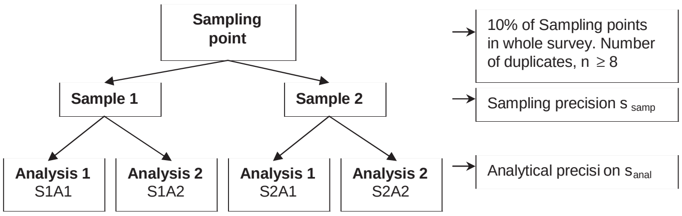

Fig 5. Balanced design for analytical duplicate analyses on each half of the sample duplicate, taken at 10 % of the sampling locations, but at not less than 8 points.

by this approach, that sampling is the main source of uncertainty in measurements, not the chemical analysis.

The bias of the analytical method was estimated as -3.3 %, by the analysis of a series of six certified soil reference materials covering the full range of Pb concentration at the site (Table 3).

The individual estimates of bias from each reference material are generally small, either in terms of absolute value at low concentration (e.g. BCR141) or when expressed as a percentage of the certified value at high concentration (e.g. NIST2710). Alternatively, the bias can be expressed as a function of concentration by performing a linear regression of the measured concentration value (c_m) on the certified values (c_{cert}), by a technique explained elsewhere (IAEA, 2003). Regression of the measured mean values on the certified values of concentration shows that the bias of the method is -3.3 %, over the entire range of concentration.

No estimate was made of the bias of the sampling method, because methods for this purpose are not yet well established.

The overall estimate of the random component of the uncertainty is 83.6 %. This means that any value of concentration for an individual location (e.g. Table 1) will be reproduced to within ± 83.6 % of the quoted value (at 95 % confidence). The bias of the analytical method can also be allowed for, either by subtracting the bias from each measurement, or by incorporating the bias into the estimation of uncertainty.

Applying this latter approach indicates that any value of concentration for an individual location has a true value that is within the range of - 80.3 % and + 86.9 % of the quoted value. This figure is the best estimate so far, but it still ignores the presence of any bias caused by the sampling procedure.

Table 2. Measurements of Pb concentration (mg/kg) on duplicated samples, taken for the estimation of measurement uncertainty at the test site. Samples were taken at the 10 coordinates described by the Sample I.D. on the regular grid shown in Table 1, using the experimental design shown in Fig 5.

Sample I.D.	S1A1Pb (mg/kg)	S1A2Pb (mg/kg)	S2A1Pb (mg/kg)	S2A2Pb (mg/kg)
A4	787	769	811	780
B7	338	327	651	563
C1	289	297	211	204
D9	662	702	238	246
E8	229	215	208	218
F7	346	374	525	520
G7	324	321	77	73
H5	56	61	116	120
I9	189	189	176	168
J5	61	61	91	119

Source: (Argyraki, 1997)

Table 3. Measured and certified Pb concentration values for certified reference materials for the estimation of the bias of the analytical method.

CRM name (n=4)	Mean (mg/kg)	Std. Dev.s (mg/kg)	Certified Value (mg/kg)	Bias (mg/kg)	Relative Bias %
NIST2709	19.7	3.2	18.9	0.8	4.23
NIST2710	5352	138	5532	-180	-3.25
NIST2711	1121.4	14.7	1162	-40.6	-3.49
BCR141	34.4	3.9	29.4	5	17.00
BCR142	36.2	4.6	37.8	-1.6	-4.23
BCR143	1297.5	33	1333	-35.5	-2.66

Source: (Argyraki, 1997)

Access to information on errors is often a limiting factor in applying this approach to contaminated land investigations. Estimates of analytical precision and bias are often made by analytical laboratories, but this information is not usually given to the customer. It is usually only used for analytical quality control within the laboratory. The customer is only told general information about the analytical method, such as 'precision was better than 10 %', 'analytical bias was generally less than 10 %' and perhaps, 'the uncertainty of measurement is generally around 10 %'. This information is inadequate for adopting the approach described here, and ignores sampling errors. Laboratories need to be encouraged to give more detailed information on precision and bias, that is specific to the customer's own sample materials.

4. FITNESS-FOR-PURPOSE:- HOW MUCH UNCERTAINTY IS ACCEPTABLE?

Once it is acknowledged that uncertainty is inevitable in investigations, the next logical question is how much uncertainty is acceptable. This is equivalent to asking 'are the measurements fit for the purpose (FFP) that is being pursued?' Chemical laboratories often set their own FFP criteria as discussed above, such as 'analytical precision must be better than 10 %'. This ignores the particular needs of the customer, as ±20 % may be good enough for some purposes. It also ignores the often dominant role of sampling precision, which laboratories usually do not estimate. Statistical techniques can be used to judge FFP criteria that address both of these issues, and can even consider the financial constraints on the system (see forthcoming CL:AIRE Technical Bulletin 8). If there is too much uncertainty in a dataset, statistics can be used to calculate how to reduce uncertainty by the required amount. For example, to reduce the uncertainty from sampling by a factor of n, it would be necessary to increase the primary sampling mass by a factor of n^2 (i.e. to reduce uncertainty in sampling by a factor of three, increase the sampling mass by a factor of nine). This can usefully by achieved by taking a composite sample at each sampling location, where each composite is made up from nine increments that are each the same mass as the original sample.

5. COMPARISON BETWEEN CONCENTRATION MEASUREMENTS AND EXTERNAL THRESHOLD VALUES

One of the main methods of interpretation in a contaminated land investigation is to compare the measured concentration values for each sampling location against some external threshold value that defines the soil as contaminated, or in need of some particular action (e.g. Soil Guideline Values, DEFRA/EA, 2002a).

As discussed in the introduction, it is possible to ignore uncertainty in this comparison. Such a 'deterministic' comparison (Fig 6a) classes any location where the measurement is lower than the threshold as 'uncontaminated', and above as 'contaminated'. If however we know the value for the uncertainty, we can see that there are occasions when, although the measurement is below the threshold value, the upper end of the uncertainty interval extends above the threshold value (Fig 6b). The true value of concentration may therefore exceed the threshold value, even though our measurement is below. Using 'probabilistic' classification, we can more reliably classify this location as 'possibly contaminated', rather than

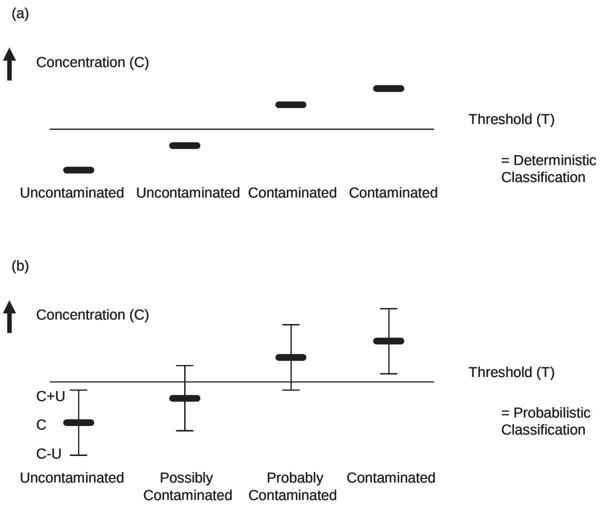

Fig 6. Comparison between (a) deterministic and (b) probabilistic classification of contaminated land, to show the effect of using estimates of uncertainty (U) to improve the interpretation of concentration measurements (C). In the case of a 'possibly contaminated' location the range of uncertainty (C-U to C+U) includes the threshold (T), showing that the true value of contaminant concentration at that location may be over the threshold value.

'uncontaminated'. Similarly, it is equally possible that although a measurement is over the threshold value, the true concentration is under this value, which we can classify this location as 'probably contaminated' (Fig 6b). Reliable decisions on classification of the land therefore require the uncertainty of measurements to be allowed for, in this probabilistic approach to classification.

Applying this probabilistic classification to the case study, we now have an estimated uncertainty value of 83.6 %, based solely upon consideration of the random error. It is possible, therefore, to classify each of the 100 locations on the site into one of these four categories (Table 4). This shows that 16 % of the locations on the site are over the threshold using the deterministic classification. A further 24 locations have concentrations over 245 mg/kg (450/1.836) which make them 'possibly contaminated'. In all therefore, nearly half the site (40 % of the locations) may have true values of Pb concentration that are possibly over the threshold, and therefore require some action to be taken. Interestingly, the uncertainty is so high that at only one location (B2) is there insignificant probability (<2.5 %) that the true concentration is under the threshold value, and the location classified as definitely 'contaminated' (i.e. measured value over 2744 mg/kg [450/(1-0.836)]).

6. CLASSIFICATION OF AN 'AVERAGING AREA' OR A WHOLE SITE

The approach to classification taken in a recent UK Government publication has begun to consider uncertainty, but on average values rather than on measurements. It bases the classification decision on a comparison of a number of sampling points within an 'averaging area' against the SGV (DEFRA/EA, 2002b). This approach does not consider the random or systematic errors in the individual measurements, but uses the variability across the set of concentration values measured across all those sampling points. Classical statistics can be used to express the standard error on the mean as $s/\sqrt{n}$, where the standard deviation (s) is divided by the square root of the number of measurements (n) used for its calculation. This shows that as n increases, the uncertainty on the mean value decreases. However, the estimate made of standard deviation also has uncertainty, expressed as a standard error on the standard deviation ($s/\sqrt{2n}$). The uncertainty on the mean value is therefore multiplied by the value of the 't' statistic that reflects this extra uncertainty for low value of n (i.e. greater than the usual value 1.96s for 95 % probability). The upper 95th percentile of the uncertainty on the mean value is therefore given by

$$US_{95} = \bar{x} + t.s/\sqrt{n}$$

If we consider the whole site for our case study, and use a value of t of 1.66 from standard tables for 95 % confidence (DEFRA/EA, 2002b) we get

$$US_{95} = \bar{x} + t.s/\sqrt{n} = 291.9 + 1.66 \times 403/\sqrt{100} = 358.8 \text{ mg/kg}$$

This value is under the SGV value for Pb of 450 mg/kg (for housing development, DEFRA/EA, 2002b) and therefore action is not warranted in the averaging area as a whole.

It is also possible to repeat this calculation for smaller averaging areas within the site.

Table 4. Measured Pb concentrations at each point on the sampling grid (mg/kg), as shown in Table 1, but classified probabilistically (Fig 6b). All values are concentrations measured on single samples except for the ten duplicated samples, for which the value given is the mean of the four values given in Table 2 for that coordinate (e.g. 787 mg/kg at coordinate A4, is the mean of the first row of Table 2). Locations with values in black are classified as 'uncontaminated' (i.e <245 mg/kg) , those in green are 'possibly contaminated' (i.e 245 - 450 mg/kg), in **bold green** are 'probably contaminated' (450 - 2744 mg/kg) and those in **bold green underlined** are 'contaminated' (>2744 mg/kg), using the 95 % confidence interval, for the estimated measurement uncertainty at this site of 83.6 %.

Row	A	B	C	D	E	F	G	H	I	J
1	**474**	287	250	338	212	**458**	**713**	125	77	168
2	378	**3590**	260	152	197	**711**	165	69	206	126
3	327	197	240	159	327	264	105	137	131	102
4	**787**	207	197	87	254	**1840**	78	102	71	107
5	395	165	188	344	314	302	284	89	87	83
6	**453**	371	155	**462**	258	245	237	173	152	83
7	72	**470**	194	83	162	441	199	326	290	164
8	71	101	108	**521**	218	327	**540**	132	258	246
9	72	188	104	**463**	**482**	228	135	285	181	146
10	89	366	**495**	**779**	60	206	56	135	137	149

For column D for example US_{95} is 468.9, indicating this area does warrant action. Considering each of the ten columns at the site, four have US_{95} values indicating action (A, B, D & F), but the other six columns do not. If the site is divided in quarters, only the top left quarter indicates action by this criterion.

There are several limitations to this approach. Primarily, it does not allow for the uncertainty in the individual measurements. To some extent the random component of the uncertainty is partially reflected in the standard deviation estimated between the different locations. The decision is made for the whole of the 'averaging area' and it is not possible to delineate particular areas for action within the area, as is possible with the first method. From the applications to this case study, it is clear that this approach is very sensitive to how the 'averaging area' is selected. It also assumes a Normal distribution of contaminant concentration, which we know is not applicable in this case study. Log-transformation of the measurements has been suggested as a possible way to achieve normality (DEFRA/EA, 2002b), although it is not entirely successful in this case (Fig 2b). The systematic error, from both chemical analysis and sampling, is not considered in this approach.

A subsequent CL:AIRE Technical Bulletin will take these ideas further to discuss the use of statistics in improving the design of sampling. It will address the questions of how many samples to take (closely related to the best sample spacing), what mass they should contain, whether to take composite samples, and if so how many increments should they contain. How can prior information be used to improve the effectiveness of sampling using a Bayesian approach? A brief overview will also be given of the techniques available for spatial mapping of contamination, including deterministic and probabilistic approaches. It will also look at techniques for judging fitness-for-purpose of measurements, including balancing the costs of site investigation and development against the inevitable uncertainties. Finally, once the uncertainty in an investigation is well known, it is possible to consider its implication for the risk assessment of the site. The uncertainty of measurement can be propagated into risk assessment to calculate the probability that an exposure limit will be exceeded, rather than making a single estimate of exposure or risk.

Acknowledgements

This bulletin was prepared by CL:AIRE staff from an original text provided by Professor Michael Ramsey at University of Sussex.

REFERENCES

- Analytical Methods Committee. 2001. AMC Technical Brief 6, Robust statistics: a method of coping with outliers (from http://www.rsc.org/lap/rsccom/amc/amc_techbriefs.htm)
- Argyraki, A. 1997. PhD Thesis, University of London (unpublished)
- BS 10175. 2001. Investigation of potentially contaminated sites. Code of practice, British Standards Institution, London
- DEFRA/EA. 2002a. SGV 10 Soil Guideline Values for Lead Contamination. Department for Environment, Food and Rural Affairs (DEFRA) and the Environment Agency (EA), London.
- DEFRA/EA. 2002b. CLR 7 Assessment of Risks to Human Health from Land Contamination: An Overview of the Development of Soil Guideline Values and Related Research. Department for Environment, Food and Rural Affairs (DEFRA) and the Environment Agency (EA), London.
- IAEA. 2003. Soil sampling for environmental contaminants, Technical Document from the International Atomic Energy Agency, Vienna (in draft form)
- Ramsey, M.H. 1998. Sampling as a source of measurement uncertainty: techniques for quantification and comparison with analytical sources. Journal of Analytical Atomic Spectrometry, 3, 97-104, (download paper, program and test data from www.rsc.org/suppdata/JA/1998/97/)

For further information contact: Professor Michael Ramsey, University of Sussex
Email: m.h.ramsey@sussex.ac.uk