Implications of data uncertainty in geo-environmental site assessment

A.G.L. Vickers
O'Connor Associates Environmental Inc., Calgary, Alberta, Canada
W.B. Kindzierski
Department of Public Health Sciences – University of Alberta, Edmonton, Alberta, Canada
T.T. Wong, J.G. Carss & J.G. Agar
O'Connor Associates Environmental Inc., Calgary, Alberta, Canada

ABSTRACT
This paper challenges geo-environmental practitioners to consider uncertainties when developing sampling plans and interpreting chemical analyses data even though environmental assessment and remediation guidelines are presented as single-valued threshold levels that define limits between "acceptable" and "unacceptable" contaminant concentrations in soil, groundwater and air. This paper also summarizes the uncertainties encountered in an environmental site assessment and discusses how analytical and sampling uncertainties can be addressed using statistical and probabilistic techniques.

RÉSUMÉ
A l'heure actuelle, les directives concernant les évaluations et remédiations environnementales ne reposent uniquement que sur une valeur seuil, définissant la limite entre une concentration "acceptable" et "non acceptable", pour un contaminant présent dans le sol, l'air ou les nappes d'eau souterraine. Cet article a pour objectif de pousser les géo-environnementalistes à cependant prendre en considération les incertitudes lors de l'interprétation de données d'analyses chimiques ou du développement d'un planning d'échantillonnage. Il présente d'une part un résumé des différentes incertitudes rencontrées lors d'une évaluation environnementale, et d'autre part comment les incertitudes de calcul et de mesure peuvent être prises en compte en utilisant des outils statistiques et probabilistes.

1 INTRODUCTION

The basic concepts behind environmental assessment can be stated as (Cotton and Emond 1981): "(1) early identification and evaluation of all potential environmental consequences of a proposed undertaking;" and "(2) decision making that both guarantees the adequacy of this process and reconciles to the greatest extent possible, the proponent's development desires with environmental protection and preservation".

These concepts are further explained in a recent Federal Court judgement (Decisions of the Federal Court 2008 FC 302) as follows:

"The Canadian Environmental Assessment Act (CEAA)… mandates early assessment of adverse environmental consequences as well as mitigation measures, coupled with the flexibility of follow-up processes capable of adapting to new information and changed circumstances. The dynamic and fluid nature of the process means that perfect certainty regarding environmental effects is not required".

The CEAA "calls for an *informed decision* by a responsible authority" and "there is a requirement to provide a rationale for its recommendations".

It would appear that, by law, some degree of uncertainty is allowed in environmental assessments; however, an environmental practitioner (from hereon, practitioner) is required to provide a rationale for the handling of uncertainties.

At present, the concept of uncertainty is rarely publicized by regulatory agencies. As a result, remediation guidelines are presented as single-valued threshold levels that define limits between "acceptable" and "unacceptable" contaminant concentrations in soil, groundwater and air. Correspondingly, practitioners commonly use a deterministic approach in environmental site assessment (ESA) and remediation. Discussions in this paper are limited to Phase II ESAs, i.e., investigations involving the collection of soil, groundwater or air samples to analyze for concentrations of potential contaminants of concern. In the following sections, uncertainties inherent in site investigations and remediation guideline development are summarized. Methods that a practitioner may use to quantify some of the uncertainties in environmental sampling are also described.

2 UNCERTAINTIES IN ENVIRONMENTAL SITE INVESTIGATIONS

The two major uncertainties in environmental site investigations are (a) uncertainty related to obtaining measurements within sampling units and (b) uncertainty associated with the variability and/or bias between sampling units (USEPA 2001). A sampling unit, also referred to as a unit of analysis, can be defined as the portion of the environmental population from which one or more samples are taken and then measured to yield test results appropriate for a specific use. The most common sampling unit would be the physical sample

taken, e.g., a 1.7 L sample of soil gas or a 600 mm long by 32 mm diameter core of soil. Figure 1 illustrates the factors contributing to total uncertainty in ESAs (after USEPA 2001).

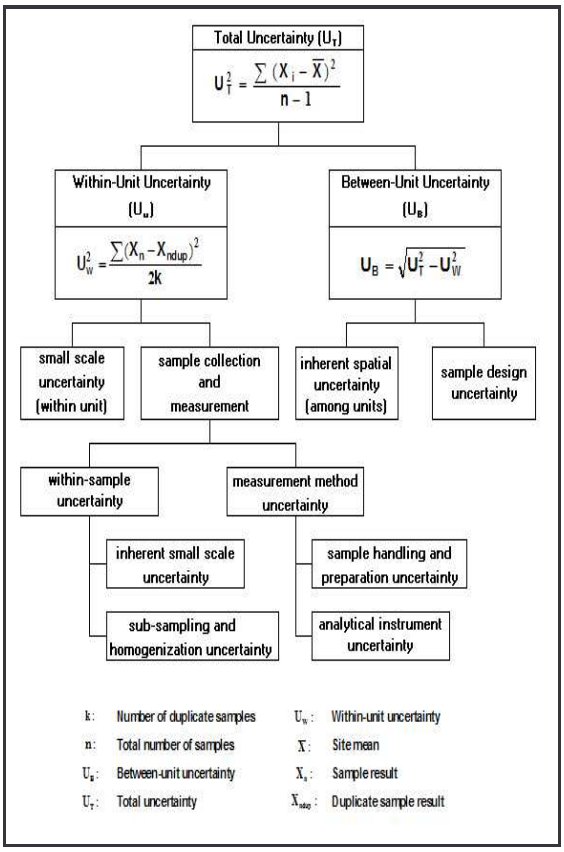


Figure 1. Components of Total Uncertainty in Environmental Site Investigations (Modified from USEPA 2001)

Between-unit uncertainty (U_B) addresses how well sampling units selected by a sampling design represent the population of interest. Within-unit uncertainty, U_W, addresses how well samples and measurements performed on them represent their true conditions. Total uncertainty, U_T, can be modelled with an additive variance formula, assuming that within-unit uncertainty is independent of between-unit uncertainty (USEPA 2001). As shown in Figure 1, within-unit variability can be estimated using field duplicate data and total uncertainty can be estimated based on the total site investigation variance. Based on the additive variance relationship, the total uncertainty, U_T, is given by:

$$U_T = \sqrt{U_B^2 + U_W^2} \qquad [1]$$

The successful management of uncertainties requires practitioners to implement quality systems, consisting of quality assurance (QA) and quality control (QC) procedures to obtain data of known and sufficient quality (Maney 2002). Statistics is a key tool for planning and determining how best to manage sources of uncertainty in environmental data sets and ensure that data based decisions are made within a desired level of confidence.

2.1 Uncertainties in Environmental Field Sampling

Many practitioners incorrectly assume that the quality of their data is primarily determined by the analytical methods used to produce the results. On the contrary, uncertainties for individual analytical measurements are often insignificant in comparison to the total statistical variation in the population (APLAC 2004). No amount of improvement in analytical precision can significantly reduce total uncertainty when the contribution of analytical uncertainty is relatively minor (Jenkins et al. 1997). Perhaps this mistake in reasoning is partially due to the relative ease in which environmental analytical laboratories quantify and mitigate their associated uncertainties. Strategies do exist for the estimation and mitigation of environmental field sampling uncertainties, and will be further discussed in Section 4.

Approximately 90% of uncertainty in environmental data can be attributed to sampling variability due to the heterogeneity of environmental matrices (Crumbling et al. 2001), encompassing both within-unit and between-unit uncertainties. Soils are more heterogeneous in comparison to other matrices such as surface water, groundwater or air and form a major source of uncertainty in sampling and sub-sampling (Jenkins et al. 1997). Factors that need to be considered in the selection and collection of samples which are representative of the population include their physical dimensions, location, timing of collection, preservation, transportation and storage (Crumbling et al. 2001).

2.2 Uncertainties in Environmental Laboratory Chemical Analyses

Environmental laboratories produce chemical analytical results, which are the product of a process involving sampling and measuring. Statistical estimation of uncertainty is the best indication of the precision (reproducibility) of an analytical test result, and most Canadian laboratories employ well-defined methods to estimate the associated method uncertainty.

In a typical laboratory, analytical method standard operating procedure (SOP) and final result calculation equation(s) (if any) are used to identify potential sources of uncertainty in the method. Uncertainty in analytical methods can arise from sources such as (CAEAL 2006): definition of the measurand; transportation, storage and handling of samples; preparation of samples and subsampling; environmental and measurement conditions; the analyst; variations in test procedures and measuring instruments; calibration standards and reference materials; software and/or methods associated with the measurement and uncertainty from the correction of the measurement results from systematic affects. Once all potential quantifiable sources of uncertainty in an

analytical method are identified, QA/QC and inter-laboratory proficiency testing (PT) data are matched with sources of uncertainty and tabulated with their associated standard deviation and/or relative standard deviation. For example, PT data are used as a source of both inter-laboratory and intra-laboratory reproducibility uncertainty. Reference samples, such as certified reference materials and laboratory control standards, are used to quantify uncertainty due to different analysts, calibration sets and standards, environmental conditions and instrument drift. Matrix spikes evaluate uncertainty due to different sample matrices; while method validation data examine uncertainty at low analyte concentration ranges, due to different analysts, instrumentation and environmental conditions. Laboratory duplicates are another measure for reproducibility uncertainty, which takes into account factors such as heterogeneity of samples, weighing, volumetric manipulations and instrument drift.

Each source of uncertainty and its associated standard deviation (SD) and/or relative standard deviation (RSD) are tabulated, eliminating doubly counted sources. Combined standard uncertainty, U_C, is then calculated using standard propagation of error roots (root sum of squares), where U_n is an individual source of uncertainty (CAEAL 2006):

$$U_C = \sqrt{\sum_n U_n^2} \qquad [2]$$

Measurement uncertainty, U_M, is expressed as an expanded U_C for a two-tailed confidence level of 95% using

$$U_M = t \times U_C \qquad [3]$$

where t is the appropriate 95% confidence level Student's t-distribution t-value. The true analytical result should then be interpreted as a value within the 95% confidence limits that are given by the reported result, X, and U_M as shown:

$$95\% \text{ confidence limits} = X \pm U_M \qquad [4]$$

The confidence level is often used loosely as a probability statement that the true result lies within the limits given by Equation 4 with a 95% probability. "Such a probability statement is, strictly speaking, inadmissible since the true result is not a random variable" (Benjamin and Cornell 1970). They further commented that:

"the engineer who has observed X and calculated confidence limits should not say 'The probability that the true mean lies between $X-tU_c$ and $X+tU_c$ is 95%' but rather just 'the 95% confidence limits on the true mean are $X-tU_c$ and $X+tU_c$'… Such 'probability statements' remain the most natural way of describing the situation and of conveying the second kind of uncertainty, that surrounding the value of a parameter. Consequently, such statements should probably not be discouraged, as they seem to express the way engineers operate within such limits…."

The use of confidence limits is most appropriate in reporting of data. It provides a convenient and concise method of reporting which is understood by a wide audience of readers, and which encourages communications of measures of uncertainty as well as simply 'best' (or point) estimates of parameters. In spite of the somewhat arbitrary nature of the conventionally used confidence levels, and in spite of the philosophical difficulties surrounding their interpretation, confidence limits remain, therefore, useful conventions." (Benjamin and Cornell 1970).

3 UNCERTAINTY IN ENVIRONMENTAL RISK ASSESSMENT AND REMEDIATION GUIDELINES

Methods currently used to develop soil and groundwater remediation guidelines in Canada (AENV 2007; CCME 2008a) do not explicitly take uncertainty into account and as a result they may be a product of overly conservative estimates of cleanup criteria by combining, through multiplication, several conservatively biased parameters. The methods are based on a coupling of traditional risk assessment process (e.g., Health Canada 2004, CCME 2008b) with an environmental fate and transport model, i.e., the Johnson and Ettinger (1991) vapour intrusion model, and are discussed further below.

3.1 Risk Assessment Process

The two types of uncertainties involved in risk assessment are (Finkel 1990): uncertainty due to variability (type A) and uncertainty due to lack of knowledge (type B). Type A uncertainty represents variability in values for a parameter, i.e., the actual distribution of values in time, space or among individuals. This type of uncertainty cannot be reduced or eliminated in risk assessment; it can only be characterized or understood. Major sources of variability in risk assessment include exposure variability and inter-individual variability in susceptibility (dose-response) (Finkel 1990). Exposure variability results from many parameters in various stages of the exposure assessment process including, but not limited to, microenvironmental and personal time-activity behaviour differences (NRC 1994). Human inter-individual variability includes differences in genetic predisposition, biological function, and behaviour (Finkel 1990).

Type B uncertainty, "true uncertainty", represents lack of knowledge about what the true value is for a parameter. This uncertainty can only be reduced by gaining knowledge to improve how much one knows about the value. The potential is very large for overly conservative estimates of remediation guidelines using standard risk assessment methods due to inadequate understanding of true uncertainty. The largest sources of true uncertainty in risk assessment occur in the dose-response assessment because of numerous assumptions and inferences which must be made. These relate to extrapolation of tested doses to estimated human doses, extrapolation between

species, and the approaches and model selections for these extrapolations. Tested doses are often values derived from animal studies and from human epidemiological studies. The process is highly uncertain because it involves extrapolating these data at relatively high exposure concentrations (in laboratory animal studies) to estimated effects on humans at much lower exposure concentrations. While the lower exposure concentrations are more often encountered in reality, health effects based on these low exposure concentrations are not generally measurable.

3.2 Guideline Development Using Environmental Fate and Transport Modelling

Environmental fate, transport, and transformation aspects of uncertainty have been separated out of the risk assessment process because of their importance in developing risk-based remediation guidelines for soil and groundwater contamination. The Johnson and Ettinger (1991) vapour intrusion (J&E) model was developed as a screening model and is no different than many other environmental fate models in use today in that it approximates various highly complex engineering and physical relationships associated with vapour migration in soil. The model uses steady state assumptions, infinite sources, limited soil biodegradation, negligible free phase, and equilibrium partitioning into air and water phases to represent advective and diffusive processes in soil. Even though the model makes some simplifying geometrical assumptions, it still requires a large number of parameters and assumptions, which translate into a potential for large uncertainties to be associated with its results. Dubus et al. (2003) offer the example that pesticide fate modelling in soils has an extensive list of parameter and assumption requirements and sources of uncertainty clearly demonstrating that the process is laced with uncertainty. Dubus et al. (2003) also importantly point out that model error and modeller preferences – which cannot be taken into account in formal sensitivity analysis – are also likely to have an important effect on the prediction of pesticide concentrations in soil.

The J&E model is admittedly complex and others have shown that complexity of models does not necessarily increase the precision of model predictions and might often decrease it (McKone 1996). "Sensitivity analysis" of model parameters in environmental modelling – which is critical to model validation – is seldom performed (Hamby 1994). In the case of the J&E model, these types of analysis are only beginning to emerge (Tillman and Weaver 2006) despite the model being introduced over 15 years ago. Tillman and Weaver (2006) further report that little published information is available on the combined effects of multiple uncertain model parameters and their effect on results in using the J&E model. Compounding conservatism by nearly a factor of 10 in model output – as represented by predicted cancer risk – was demonstrated by Tillman and Weaver (2006) during formal sensitivity analysis using a multiple-parameter uncertainty approach versus a single-parameter (one-parameter-at-a-time) uncertainty approach.

Those involved in development of remediation guidelines all too often fail to fully consider problems when there are data limitations or other issues that introduce uncertainties. The importance of thorough uncertainty analysis would mostly be acknowledged, but all too often uncertainties are not specifically quantified or understood in favour of qualitative statements that conservatism in the guideline development process counter-balances uncertainty in the data. In all but the most sophisticated and costly risk-based processes for guideline development, uncertainty analysis is seldom conducted — usually at the end of the process and often only in a qualitative way. Unfortunately, just using high (conservative) end value for each parameter endpoint in the guideline development calculation results in conservatism compounded to an extent that is not understood well enough. The soil quality guideline development processes that currently exist (e.g., AENV 2007; CCME 2008a) can be viewed as generally cautious approaches, for which reasonable data and sensible knowledge of uncertainties (i.e., established through some type of formal sensitivity analysis) are likely to yield reasonable estimates that will protect public health. However, where poor data and lack of knowledge contribute to uncertainty, and limited attention is paid to analyzing uncertainty in the development process, resulting guidelines are very likely to be over-protective. For this latter case, the ability of soil and groundwater contamination levels that are a factor of two or possibly more than current guideline levels to pose health risks should be interpreted with a great deal of caution and healthy scepticism.

4 DEALING WITH UNCERTAINTIES IN ENVIRONMENTAL SITE INVESTIGATIONS

As discussed in the previous sections, uncertainties in environmental site investigations are incurred in: field sampling, chemical analysis, the environmental risk assessment process and in guideline development. Uncertainties in the risk assessment process and guideline development are the responsibility of regulatory agencies. Some of the methods that can be used by practitioners to quantify, interpret and mitigate uncertainties in environmental site investigations are outlined in the following.

4.1 Quantifying Uncertainties in Sample Collection and Sampling Plan Development

Environmental site investigations generally involve collecting samples of soil, water and/or air media, and measuring the concentrations of the potential contaminants of concern (pcoc). Traditionally, sampling plans for these investigations have been developed in a subjective manner by practitioners. The selection of sampling locations is dependent on a number of factors such as the experience level and personal judgement of the project team members, the available historical land

use information, budget, client preferences or third party concerns. It is recognized by most experienced practitioners that successful identification of the pcoc and the likely impacted areas depends on the accuracy and level of detail available regarding historical land uses. The quality of historical information about contaminated sites is variable and, in some cases, critical historical land use details have been lost. The absence or unavailability of valuable information has the potential to unknowingly minimize the value of the professional judgement. Statistics Canada (2006) classifies judgement sampling, along with convenience or haphazard sampling (where samples are collected based solely on accessibility), as non-probability sampling methods in which "it is impossible either to estimate sampling variability or to identify possible bias" (Statistics Canada 2006).

As discussed in Section 2, the total uncertainty in a site investigation is made up of within-unit and between-unit uncertainties. Within-unit uncertainties in soil sampling can be managed by using "correct sampling procedures" and equipment, as detailed by Gy's sampling theory (GST) for particulate matter (Gerlach et al. 2002; Gerlach and Nocerino 2003). GST stresses the importance of carefully considering the processes used to physically obtain and extract samples in order to achieve a representative, non-distorted sample from a sampling unit. The amount and weight of a sample should be considered in relation to the matrix particle size and shape (Gerlach et al. 2002; Gerlach and Nocerino 2003). One limitation to GST is that it is not intended for volatile and semi-volatile constituents (Gerlach and Nocerino, 2003). GST often involves drying and communition (particle size reduction), which could result in significant losses of volatile constituents (ASTM, 2003).

Between-unit uncertainties can be quantified and managed using a probability-based sampling design, which determines the required type, quality, quantity and placement of samples to ensure that the resulting data are statistically representative of the sampling units and overall population of interest, and can support a sufficiently confident decision (Pulsipher et al. 2003). Visual Sampling Plan (VSP) software (PNNL 2008) is a tool that can help manage between-unit uncertainties by creating a probability based sampling design. VSP can be used to assess the relative contributions of sampling and analytical uncertainties to the total uncertainty, and evaluate how to best to reduce uncertainty. Methods to reduce uncertainty include: obtaining more samples and/or improving sampling technique or conducting replicate analyses and/or improving precision of the analytical method. Another strategy for managing between-unit uncertainties is the use of composite sampling. Composite sampling can help to improve the representativeness of samples, and is a more cost-effective strategy for reducing spatial uncertainty than merely obtaining a greater number of samples (Jenkins et al. 1997). However, composite sampling is not an appropriate method for certain volatile and semi-volatile contaminants due to concerns regarding constituent loss during compositing procedures. Moreover, when using composite sampling, the comparative guideline value may have to be divided by the number of samples used in compositing (Patil 1995). Soils are more heterogeneous in comparison to other matrices such as surface water, groundwater or air; soil heterogeneity is a major source of uncertainty in sampling and sub-sampling (Jenkins et al. 1997). Factors that need to be considered in the selection and collection of samples that are representative of the population include their physical dimensions, location, timing of collection, preservation, transportation and storage (Crumbing et al. 2001)

A probability-based sampling plan supports valid inference of the mean and variance of the target population, provides quantitative estimates of uncertainty and variance and indicates limits on uncertainty associated with a decision. The idea of applying statistical and probability concepts to improve sampling plan development and decision-making processes has been around for more than 50 years (e.g., Freeman et al. 1948; Gilbert, 1982). Details of statistical methods for environmental site sampling are well presented in Gilbert (1987).

Since the late 1980s, the US Environmental Protection Agency (USEPA) has been advocating the use of statistical evaluation methods for establishing environmental data quality objectives (DQO) and for assessing, in a statistically unbiased manner, whether the number and spatial distribution of the samples are sufficient for characterizing the pcoc at a site with an acceptable level of confidence. However, the use of a probability-based sampling plan without careful consideration of site-specific factors and economics can be hampered by: the appearance that the sampling locations do not make sense, by mathematical concepts that are difficult to explain to the general public, and by little perceived control of sampling costs. To facilitate the implementation of a significant portion of a DQO program, the Visual Sampling Plan (VSP) software (PNNL 2008) has been developed to design statistics-based soil and sediment sampling plans for impacts due to metals or unexploded ordnances. The application of VSP to establish the required areal extent of a remedial excavation is described as follows.

4.2 Example: Development of Sampling Plans Using VSP

Sampling methods allowed for in VSP include: simple random sampling, systematic grid sampling, stratified sampling, cluster sampling, sequential sampling, collaborative sampling, ranked set sampling and judgment sampling (not probability-based) (Matzke et al. 2007). In the following, results using systematic grid sampling to determine, with a specified probability, the locations of hot-spots of a specified size and shape in a study area are discussed.

By definition, a hot-spot is "a local contiguous area that has concentrations that exceed a threshold value" (Matzke et al. 2007). In a site assessment, the size and shape of hot-spots are determined through consultation among regulators, practitioners and, sometimes, stake-holders. In this example, the hot-spots are arbitrarily chosen to be circular and 1.0 m in radius and

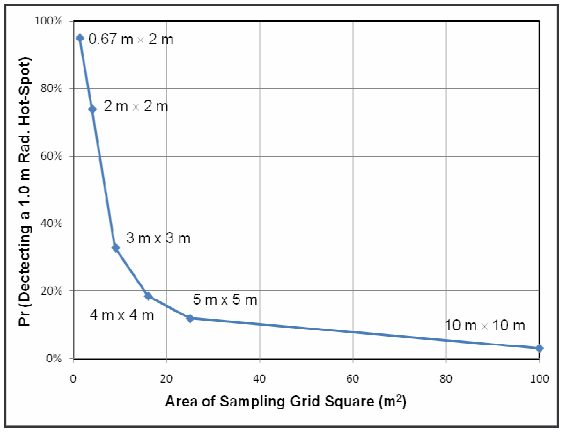


Figure 2. Decreasing Probability of Detection with Increasing Grid Size

systematic grid sampling is used. Figure 2 shows a significant reduction in the probability of detection, p(detect), with increasing sampling grid spacing using VSP. For example, 0.67 m x 2 m grid (area = 1.34 m^2) is able to detect a 1.0 m radius hot-spot with a p(detect) = 95%; whereas a 3 m x 3 m grid (area = 9.0 m^2) can only achieve a p(detect) of 33%.

When locating a hot-spot, VSP calculates the necessary sampling grid size based on input hot-spot size and β, the probability of not detecting a hot-spot or the "consumer's risk". If the probability of detecting at least a hot-spot is p(detect), then β = 1 – p(detect). The commonly accepted values of β are between 5% and 10%. For a p(detect) = 95%, β = 5%. Gilbert (1987, Chapter 10) presented three graphs relating β to the spacing required for regular square, rectangular and equilateral triangular grids, respectively. However, the sampling grid spacing thus obtained can be increased if prior information about the probability that a hot-spot actually occurs is available. This *a priori* probability, p(A), can be obtained using pilot surveys or other knowledge such as historical data. The procedure of adjustment is based on conditional probability to determine the probability that a hot-spot exists and is not detected, and is described in detail by Gilbert (1987, p.128).

For example, if the acceptable probability (β) for a sampling grid to miss a circular hot-spot 1 m in radius is 5%, the spacing for a square grid would be 1.69 m without knowing p(A). If p(A) = 15%, then modified consumer's risk is calculated to be 0.298 (Gilbert 1987,

Table 1. Modification of (Square) Grid Spacing by *A Priori* Probability of Exceedance for a Consumer's Risk of 5%

	A Priori Probability p(A)				
	10%	15%	20%	25%	50%
Grid Spacing (m)	2.45	2.15	2.03	1.95	1.69

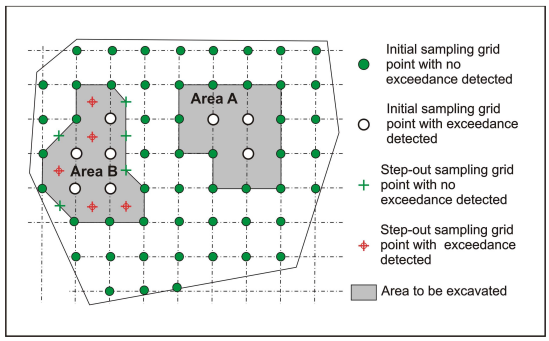


Figure 3, Schematic Showing Delineation Procedure for a Remedial Excavation

Equation 10.5) and the grid spacing obtained from Gilbert's chart is 2.15 m. The variation of (square) grid spacing with p(A) is summarized in Table 1 for a consumer's risk of 5%.

To allow for a highly non-uniform distribution of contaminants across a site, the sampling grid spacing may also be locally refined within an initial grid using a step-out procedure. Figure 3 shows a 3 m x 3 m grid used to delineate an area to be excavated in order to remediate metal impacts. As shown in Area A, the perimeter of excavation is determined by bounding the grid points where exceedances are detected with a line connecting the grid points where guideline exceedances are not detected. To ascertain (as well as to increase) p(detect), additional samples can be taken at step-out locations, i.e. from grid points with exceedance to the centroids of the adjacent grid squares as shown on Figure 3. The step-out grid will thus be at $3/\sqrt{2}$ = 2.12 m spacing. Depending on the chemical analyses results, the step-out procedure may reduce the area to be excavated. For example, without stepping out, Area B as shown in Figure 3 will cover 11 grid squares (99 m^2). Based on the chemical analyses results at the step-out sampling points, the area to be excavated is reduced to 9 grid squares (81 m^2). The economy of reducing the extent of excavation would have to be balanced against the increased cost of sampling and chemical analyses.

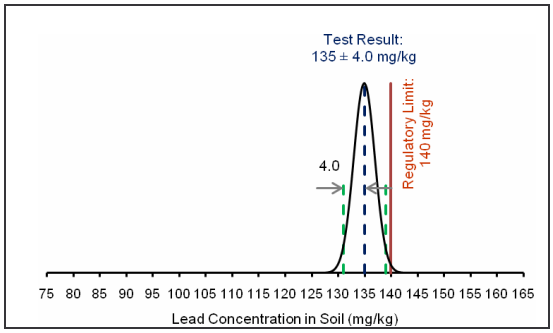
The above discussion does not consider possible vertical variation in chemical concentrations. To account for vertical variations, samples could be taken at depth-specific intervals. The final extent of excavation is determined by overlaying the areal extents obtained for each depth-specific layer.

4.3 Interpreting Laboratory Data Uncertainty

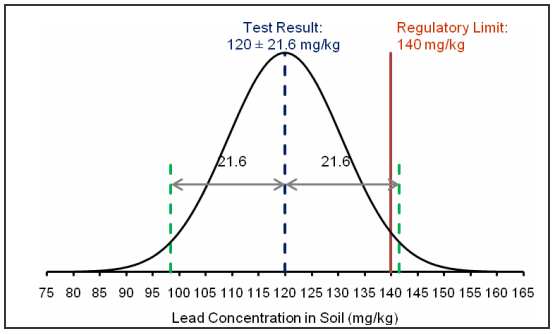
In 2003, recognizing that measurement uncertainties are inherent in chemical analyses, the Canadian Association for Environmental Analytical Laboratories (CAEAL) began requiring laboratories accredited under ISO/IEC 17025 to provide information on measurement uncertainty in test reports when requested by clients or when the uncertainty affects compliance to a specification limit (CAEAL 2006; ISO/IEC 2005). Unfortunately, the significance of uncertainty has not

yet been fully appreciated by many Canadian practitioners, who continue to compare laboratory results with a single-valued regulatory guideline in a deterministic manner.

How would a practitioner responsibly compare and report laboratory results, complete with measurement uncertainties, with a single regulatory standard for decision making? The importance of the laboratory measurement uncertainty can be illustrated by the following example. Consider laboratory test results as shown in Figure 4. Two test results with their corresponding uncertainties are modelled as normal distributions in Figure 4(a) and Figure 4(b), respectively. As shown in Figure 4(a), the test result was measured at 135 ± 4 mg/kg. Statistically, 95% confidence intervals will contain the true mean 95% of the time. In this example, one particular 95% confidence interval is obtained as 135 ± 4 mg/kg and this interval does not contain the regulatory limit, which is set at 140 mg/kg. As discussed in Section 2.2, a practitioner can loosely interpret that the test result lies between 135 ± 4 mg/kg with a probability of 95% and conclude that there is less than 5% chance that the true sample result may exceed the regulatory guideline. In Figure 4(b), the regulatory limit falls within the interval 120 ± 21.6 mg/kg. In this case, the practitioner cannot draw the same conclusion, regardless of the fact that the actual reported test result is considerably lower than the criterion as shown in Figure 4(a).



a) Test result plus uncertainty does not exceed the regulatory limit



b) Test result plus uncertainty exceeds the regulatory limit

Figure 4. Comparing Test Results with Uncertainty to Regulatory Limit

5 CONCLUSIONS

An environmental practitioner encounters uncertainties in every step of a Phase II Environmental Site Assessment including sampling, chemical analysis and evaluating analytical results against regulatory guidelines. It is the responsibility of professional environmental assessment practitioners to state their rationale when dealing with these uncertainties.

As discussed above, some of these uncertainties can be quantified using probability-based sampling plans and statistics. However, substantial uncertainty can also exist in development of guidelines – which are used by practitioners to interpret site investigation results and make decisions respecting the need for additional investigation or remediation. Unfortunately, uncertainties associated with guideline development are seldom adequately quantified or clearly communicated by the Canadian regulatory agencies.

ACKNOWLEDGEMENTS

The writers would like to thank Ms. Aurore Kurc for providing the French translation of the abstract of this paper.

REFERENCES

AENV (Alberta Environment). 2007. *Alberta Tier 1 Soil and Groundwater Remediation Guidelines*. Alberta Environment, Edmonton, AB. June 2007.

APLAC (Asian Pacific Laboratory Accreditation Cooperation) Policy, 2004. *Interpretation and Guidance on the Estimation of Uncertainty of Measurement in Testing,* Asia-Pacific Laboratory Cooperation, (APLAC), North Melbourne, Australia. March 2004.

ASTM (American Society for Testing and Materials). 2003. *Standard Guide for Laboratory Subsampling of Media Related to Waste Management Activities.* ASTM Standard D6323, ASTM International, West Conshohocken, PA.

Benjamin, J.R. and Cornell, C.A. 1970. *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill Book Company, New York.

CAEAL (Canadian Association for Environmental Analytical Laboratories). 2006. *CAEAL Policy on the Estimation of Uncertainty of Measurement in Environmental Testing.* PT19-CAEAL Policy on Uncertainty, Rev 1.8. Viewed 2007-09-17. http://www.caeal.ca/P19_CAEAL_Unce_Pol.pdf

CCME (Canadian Council Ministers of the Environment), 2008a. *Canada-Wide Standards for Petroleum Hydrocarbons in Soil: Technical Supplement.* CCME, Winnipeg, MB. January 2008.

CCME (Canadian Council Ministers of the Environment), 2008b. *Canada-Wide Standards for Petroleum Hydrocarbons in Soil: Scientific Rationale.* CCME, Winnipeg, MB. January 2008.

Cotton, R. and Emond, D.P. 1981. Environment impact assessment in *Environmental Rights in Canada*, J. Swaigen, ed. Quoted in Decisions of the Federal Court, 2008.

Crumbling, S.M., Groenjes, C., Lessnik, B., Shockley, J., van Ee, J., Howe, R., Keith, L. and McKenna, J. 2001 Managing Uncertainty in Environmental Decisions. *Environmental Science and Technology*, October, p. 405A - 409A.

Decisions of the Federal Court. 2008. *Pembina Institute for Appropriate Development v Canada (Attorney General) 2008 FC 302*. Viewed 2008-04-21. http://decisions.fct-cf.gc.ca/en/2008/2008fc302/2008fc302.html.

Dubus, I.G., Brown, C.D. and Beulke, S. 2003. Sources of uncertainty in pesticide fate modelling. *Science of the Total Environment*, **317**: 53-72.

Finkel. A. 1990. *Confronting Uncertainty in Risk Management*. Center for Risk Management Resources for the Future, Washington DC.

Freeman, H.A., Friedman, M., Mosteller, F. and Willis, H.A. 1948. Sampling Inspection: Principles, Procedures and Tables for Single, Double and Sequential Sampling. In *Acceptance Inspection and Quality Control Based on Percent Defective*, McGraw-Hill Book Company, Inc., New York, USA, 395 pp.

Gerlach, R.W. and Nocerino, J.M. 2003. *Guidance for Obtaining Representative Laboratory Analytical Subsamples from Particulate Laboratory Samples.* EPA/600/R-03/027, U.S. Environmental Protection Agency, Viewed September 17, 2007. http://www.epa.gov/esd/tsc/images/particulate.pdf

Gerlach, R.W., Dobb, D.E., Raab, G.A., Nocerino, J.N. 2002. Gy Sampling Theory in Environmental Studies, I: Assessing Soil Splitting Protocols. *Journal of Chemometrics,* **16** (7): 871-878.

Gilbert, R.O. 1982. Some statistical aspects of finding hot spots and buried radioactivity. In *TRAN-STAT Statistics for Environmental Studies*, No. 19, PNL-SA-10274, Pacific Norwest National Laboratory, Richland, Washington, USA.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. John Wiley & Sons, Inc., New York, USA, 320 pp.

Hamby, D.M. 1994. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, **32**: 135-154.

Health Canada. 2004. *Federal Contaminated Sites Guidance on Human Health Risk Assessment in Canada*. Health Canada, Ottawa, ON.

ISO/IEC (The International Organization for Standardization and the Electrotechnical Commission) 2005. *General Requirements for the Competence of Testing and Calibration Laboratories*. ISO/IEC 17025, ISO Committee on conformity assessment (CASCO).

Jenkins, T. F., Grant, C.L., Brar, G. S., Thorne, P.G., Schumacher, P.W., and Ranney, T.A. 1997. Sampling Error Associated with Collection and Analysis of Soil Samples at TNT-Contaminated Sites. *Field Analytical Chemistry and Technology*, **1**(3): 151-167.

Johnson, P.C. and Ettinger, R.A. 1991. Heuristic model for predicting the intrusion rate of contaminant vapors into buildings. *Environmental Science and Technology*, **25**: 1445-1452.

Maney, J.P. 2002. Optimizing Data Collection Design. *Environmental Science and Technology*, **36**(19): 383A - 389A.

Matzke, B.D., Wilson, J.E., Nuffer, L.L., Dowson, S.T., Gilbert, R.O. and 7 others 2007. *Visual Sampling Plan Version 5.0 User's Guide.* Pacific Northwest National Laboratory, Richland, WA, USA.

McKone, T.E. 1996. Alternative modelling approaches for contaminant fate in soils: uncertainty, variability, and reliability. *Reliability Engineering and System Safety*, **54**: 165-181.

NRC (National Research Council). 1994. *Science and Judgment in Risk Assessment*. National Academy Press, Washington, DC.

PNNL (Pacific Northwest National Laboratory) 2008. Visual Sampling Plan (VSP). Viewed 2008-04-25. http://vsp.pnl.gov/.

Patil, G.P. 1995. Composite Sampling, *USEPA Observational Economy Series, Volume 1,* EPA 230-R95-005.

Pulsipher, B., Gilbert, R. and Wilson, J. 2003. *Measurement Uncertainty in Visual Sampling Plan (VSP)*. PNNL-SA-38977, National Environmental Monitoring Conference, July 2003. Viewed 17 April 2008; http://dqo.pnl.gov/vsp/PNNLSA38977.pdf

Statistics Canada 2006. Statistics: Power from Data! Sampling Methods – Probability Sampling. Viewed 2008-04-24. http://www.statcan.ca/english/edu/power/ch13/probability/probability.htm.

Tillman Jr., F.D. and Weaver, J.W. 2006. Uncertainty from synergistic effects of multiple parameters in the Johnson and Ettinger (1991) vapor intrusion model. *Atmospheric Environment*, **40**: 4098-4112.

USEPA (United Sates Environmental Protection Agency). 2001. *Guidance on Data Quality Indicators*. EPA QA/G-5i Peer Review Draft, Office of Environmental Information, Washington, DC.