The Triad Approach to Managing the Uncertainty in Environmental Data

Deana M. Crumbling
U.S. Environmental Protection Agency, Office of Superfund Remediation and Technology Innovation, 1200 Pennsylvania Ave., NW, Washington, DC 20460 USA
E-mail: crumbling.deana@epa.gov; Phone: 703-603-0643

INTRODUCTION

The first-generation data quality model that equated environmental data quality with analytical quality was a useful starting point for the site restoration community. However, for many contaminated site projects this model for establishing reliable data fails in practice because it does not consider numerous relevant variables. The inability of the "analytical quality = data quality" model to ensure data representative of the intended decision is an important reason why cleanup projects tend to take years of repeated site characterization efforts to get to closure or successful remediation. Cost-effective, efficient, and defensible cleanup projects depend on an environmental data quality model that explicitly includes all major contributors to data uncertainty.

The U.S. EPA has articulated the Triad approach as a practical framework that synthesizes new technologies and advancing science with evolving regulatory and engineering practices governing site cleanup. The Triad approach rests on the foundation of managing decision uncertainty. Managing data uncertainty, especially sampling uncertainties, is critical when decisions will be based on data. The second-generation framework offered by the Triad approach not only increases decision confidence, but also decreases project lifecycle costs by evolving the site conceptual model in real-time (using dynamic work strategies) whenever feasible. Projects implemented using Triad principles typically show lifecycle cost savings in the neighborhood of 30-50% as compared to first-generation strategies for site work. A key reason for Triad cost-savings is that characterization is performed very efficiently and accurately, avoiding decision errors that waste resources.

The purpose of site investigation and characterization is to develop an understanding of the nature and extent of contamination that is accurate enough to support correct decisions—whatever those decisions may be. The most important decision early in a project may simply be: Is there contamination present in quantities that could pose a risk to receptors such that more in-depth investigation is required? If the answer is yes, a more accurate conceptualization of contaminant mass, distribution, and mobility will need to be developed to support subsequent decisions about exposure risk and risk mitigation. If the early decision is faulty because isolated data points give misleading information, two types of decision errors are possible: 1) resources spent needlessly characterizing insubstantial contamination; or 2) unacceptable exposure risk from significant contamination that was missed by the sampling program.

Contamination that is uniformly distributed throughout a matrix is easy to detect if it is present, and easy to conclude it is not present if isolated samples do not detect it. However, the physical mechanisms by which pollutant release and migration occur ensure that contaminants are rarely spread evenly throughout a site's boundaries. As illustrated later in this paper, contaminant heterogeneity easily lead to both kinds of decision errors unless the decision maker develops and tests predictions about where contamination would be if present. Heterogeneity can also produce misleading pictures of contamination if data uncertainties are not controlled. The model that predicts and describes contaminant nature and extent is termed the "conceptual site model" or CSM. It is the mental picture on which decision maker ultimately bases all project decisions. Consensus among stakeholders and other involved parties is possible only when all are confident that the final CSM accurately represents site contamination.

THE CONCEPTUAL SITE MODEL

The CSM is a picture of site contamination. Building a CSM begins with the "story" of

- how contamination was released and what mechanisms cause migration or transformation;
- what distinct spatial patterns or contaminant distributions are created by mechanisms of release, fate, and transport;
- what receptors might be exposed to contamination and to how much; and
- what might be done to cost-effectively and efficiently mitigate potential exposures.

The preliminary or initial CSM is built (i.e., predicted) from

- information gleaned from the site history,
- knowledge of how contaminants are typically released,
- knowledge of how they behave once released to the environment, and
- existing site data, not just for contaminant concentrations, but also for parameters that influence contaminant behavior (e.g., pH, organic carbon content, particle size, porosity, stratigraphy, topology, etc.).

The preliminary CSM functions as the working hypothesis about site contamination that will be continually tested and refined as more information (including data) are integrated into the contamination model. The more closely the CSM depicts reality with respect to the intended decisions, the more cost-effective and successful those decisions can be. The more the model deviates from actual site conditions, the more likely that risk decisions and remedial designs will be incorrect. The CSM guides design of sampling and analysis plans to fill data gaps obstructing confident decision-making. The CSM is the tool used to

- predict the degree of contaminant heterogeneity and the nature of spatial patterning;
- verify whether those predictions were accurate;
- assess whether heterogeneity can compromise the performance of statistical sampling plans;
- understand "data representativeness;"
- communicate a common understanding and vision of the project among all stakeholders, and
- integrate knowledge of heterogeneity and spatial patterning into decisions about exposure pathways, remedy selection, treatment system design, and strategies for long-term monitoring; i.e., reconcile the scale of sample collection/data generation to the very different scale(s) at which project decision(s) are made.

An important function of the CSM is to identify and delineate different contaminant populations. Contaminant release and migration mechanisms typically create spatially distinct populations where impacted media are interspersed among non-impacted media. This intermingling of populations can occur on macro (between-sample scales) and micro (within-sample) scales. Both can have severe repercussions on the ability of contaminant concentration results to reliably represent contaminant nature and extent in support of project decisions. Contaminants may migrate through narrow flow channels (termed preferential pathways) whose small spatial volumes are hard to detect, but may be a primary exposure route.

Knowledge of the physical mechanisms of contaminant release and migration can be used to predict contaminant locations and the degree of spatial patterning. These predictions form the basis for drawing up the preliminary CSM (or perhaps two or three competing preliminary CSMs), which are then tested as data collection confirms, rejects, or modifies the current CSM. Populations are most productively defined by combining knowledge of spatial patterning with potential site decisions. For example, Figure 1 depicts a wind deposition scenario creating surface soil contamination in a pattern of coarse concentration contours that span five orders of magnitude.

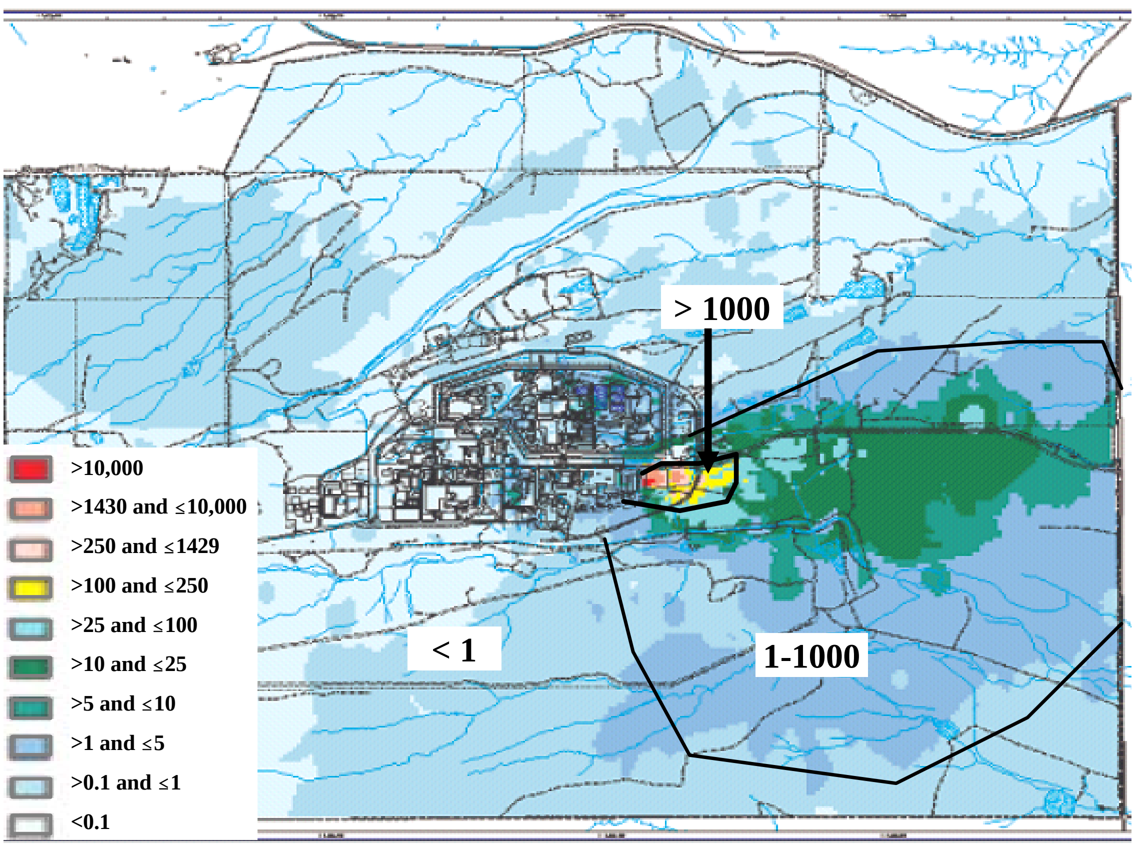

Figure 1. Surface contamination pattern caused by atmospheric deposition as influenced by regional wind patterns.

Obviously, a high sampling density (and a large budget) is required to achieve delineation at the fine scale depicted in Figure 1. A fine scale may not always be needed to effectively manage a site. Target populations can be defined using the project decision framework to determine the scale required for delineation. By way of illustration, a hypothetical scenario depicted in Figure 1 might require delineation of just three populations to support decisions about contaminated soil removal to support land reuse: natural background (up to 1, for which no action is required), between 1 and 1000 (for which landfill disposal is the likely remedial option), and greater than 1000 (destructive treatment is required). Efficient characterization is possible only if the decision framework is understood *before* the sampling and analysis plan is designed: a one-size-fits-all sampling plan will not work.

"Sampling uncertainty" occurs because environmental matrices are heterogeneous in both physical composition and in pollutant distribution. The term embraces a number of factors that introduce variability into analytical results. Analytical data can be misleading when sampling variables are not controlled. Decision errors occur when accurate analytical results generated from tiny samples are assumed by data users to represent the concentrations of much larger volumes of matrix, but that extrapolation is invalid because confounding variables have not been acknowledged or controlled. Figure 2 illustrates how unjustified extrapolation of analytical results to larger volumes of matrix can produce inaccurate CSMs that lead to faulty decisions.

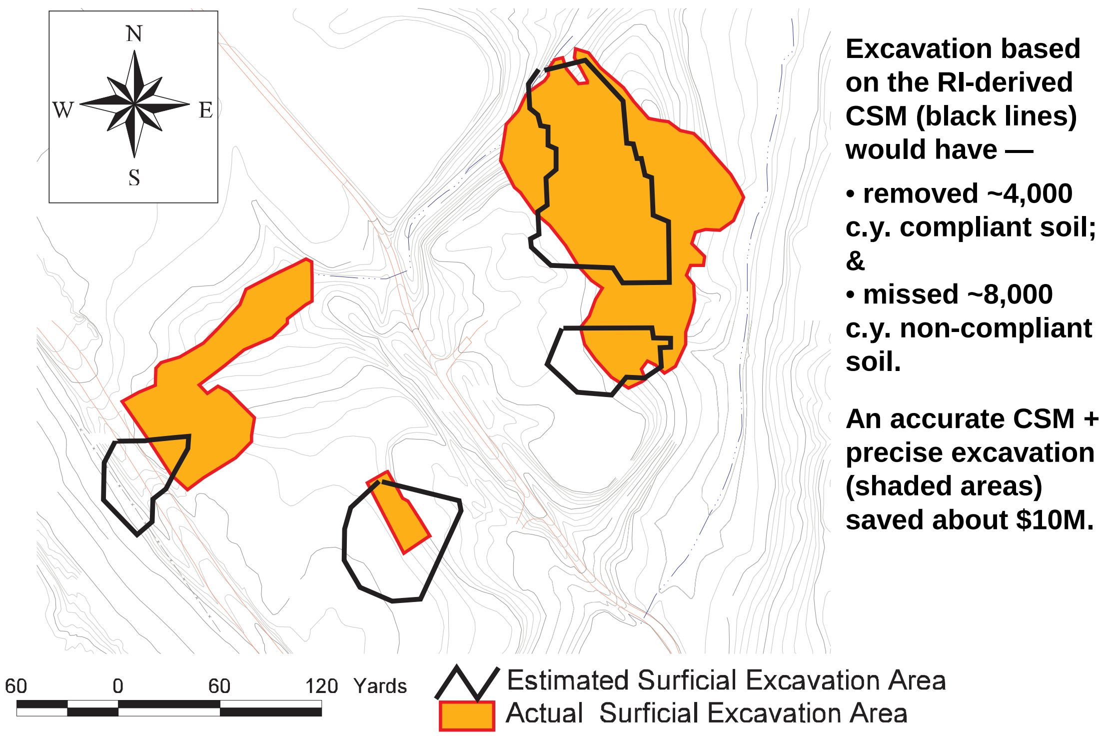

Figure adapted from Argonne, 2002

Figure 2. An inaccurate CSM can lead to costly decision errors.

The CSM portrayed by the black outline predicts the extent of contaminated surficial **soils** requiring removal based on data collected using a traditional Remedial Investigation (RI) approach of sampling with fixed laboratory analysis. Before cleanup could be implemented, the team became concerned about excessive uncertainty in the bounded areas. Taking data uncertainty into account, the volume of soil needing removal and disposal (at $300 per cubic yard) ranged could range as low as 3,000 or as high as 46,000 cu. yd. Confident remedial planning based on the RI data was impossible, but newer technologies were available to provide high density, real-time data that could manage the decision uncertainty. The team decided to implement an adaptive sampling and analysis program that was integrated in real-time with soil removal activities. By the end of the cleanup, the actual (very high confidence) CSM for surficial soil contamination was demonstrated to be the shaded areas. The total volume removed (both surficial and deeper layers) was 45,000 cu. yd. Post-cleanup sampling confirmed that on-site cleanup goals were attained. Pre-disposal testing of waste soil confirmed the "dirty" status of removed soil. Under a Triad approach, $200,000 was spent to re-characterize the site to manage both decision uncertainties. If the CSM predicted by traditional sampling and analysis had been followed, over $1.5 million would have been wasted just to needlessly remove and dispose of clean soil. Since post-remediation sampling would have discovered that 8,000 cu. yd. of "dirty" soil were missed, one more repeat cycle of characterization and removal would have been required (assuming an accurate CSM was achieved the second time). By breaking the characterize poorly—remediate poorly—recharacterize cycle, a $200,000 investment yielded an estimated savings of $10 million in cost-avoidance (DOE, 2001).

Studies with modern tools show that heterogeneity impacts **groundwater** sampling as well. There is now ample evidence that vertical stratification of common pollutants occurs in many lithological settings. The concentration of contaminants can change drastically over short depth intervals. For example, chlorinated volatile organic compounds (VOCs) concentrations were observed to change 2,500 µg/L over a vertical distance of 3.4 feet in one well, and from 7,300 to 17,500 µg/L over a vertical distance of 5 feet in another well (Vroblesky and Peters, 2000).

When well screens span different populations, purging and sampling the well can cause uncontrolled mixing between distinct populations, creating intermediate data results that produce erroneous CSMs. This is illustrated by Figure 3, which shows the results of a U.S. Geological Survey (USGS) study comparing sampling techniques for wells with long screens (Huffman, 2002). Chlorinated VOCs were analyzed by the same analytical method on water samples collected in two different ways: traditional low-flow purging with a submersible pump (left-hand panel) versus passive diffusion bag samplers (PDBs, right-hand panel). PDBs consist of a semi-permeable polyethylene "baggie" filled with distilled water that is lowered into a groundwater well. The PDB remains undisturbed in the well for 2 to 3 weeks, which allows certain contaminants to pass through the bag into the distilled water. After equilibration, the sampler is removed from the well and emptied into traditional vials for submittal to analysis. Figure 3 compares the two different sampling techniques for the same well field for trichloroethane (TCA) results. It is clearly evident that vertical stratification exists in wells 6-S-21 and MW-7. In well 6-S-21, mixing at the population boundary by the traditional sampling technique created an intermediate result. The PDBs preserved information about distinct contaminant populations, producing a different, yet more accurate CSM to guide decisions about contaminant extent and remediation.

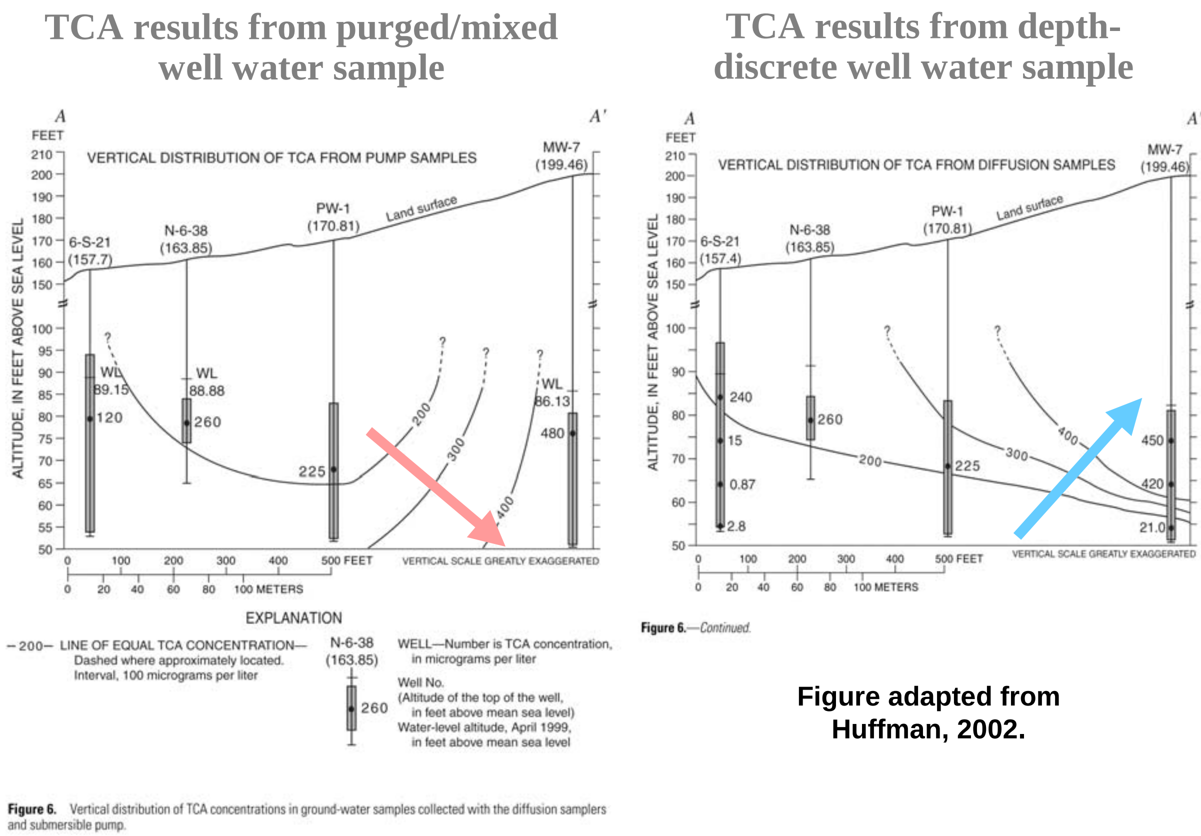

Figure 3. Sampling the same well field in different ways produces different CSMs.

THE CHALLENGE OF DATA REPRESENTATIVENESS

Generating "representative" data is not a simple matter when heterogeneous environmental matrices are involved. Figure 4 introduces the range of variables that have been found to impact the ability of data to provide reliable information for decision-making purposes. Variables that contribute to the data uncertainty can be coarsely grouped into three categories. The length of this paper limits discussion to only one variable, but a very important one regularly neglected by the environmental community. Yet each variable forms a link in the data quality chain, and each link must be intact if data are to be representative of the intended decision. The first step for ensuring representative data is to understand exactly how the data will be used in the decision-making process. The intended decision will define what population should be targeted by data collection and analysis. Sampling and analytical procedures must be tailored to the target population to avoid a common cause of data uncertainty: uncontrolled mingling of different populations.

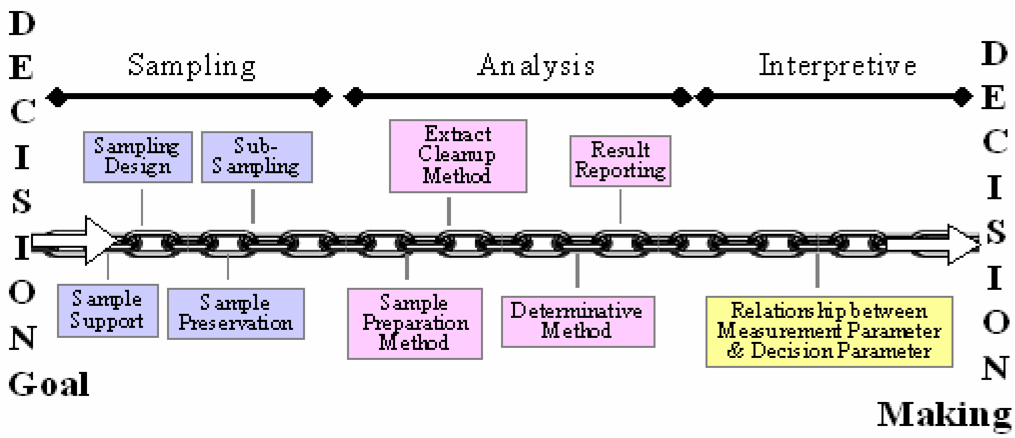

Figure 4: Variables that contribute to data representativeness.

Since contaminated sites typically encompass two or more contaminant populations, no facet of data collection and analysis can be left to chance. Each variable must be selected to maintain the chain of "data representativeness." Breaking that chain can produce data that misleads decision-makers into erroneous conclusions and actions.

SAMPLE SUPPORT: A CRITICAL VARIABLE FOR REPRESENTATIVE DATA

The term "sample support" is unfamiliar to the environmental field, yet the term was introduced to the cleanup community in several EPA documents of the early the 1990s. The term even appeared in a widely circulated U.S. EPA Superfund guidance (EPA, 1993, p. 41), but the concept never caught on. The term comes from statistics language to collectively describe the physical attributes of a specimen that help determine what the analytical result will be. These attributes apply both to samples taken from the parent matrix in the field and to subsamples taken from jars in a laboratory. For environmental samples, they commonly include 1) the mass/volume of the sample or subsample; 2) the spatial orientation/dimensions of the sample collection device which helps determine the spatial dimensions of the sample (for example, visualize a long thin corer versus a flat-bottomed scoop; and 3) particle size. Differences in sample support can cause analytical results to be different, independent of any variability in the analytical method itself. The reason is that these attributes help define different contaminant populations. Sample support is listed in Figure 4 as the first variable in the second-generation data quality because it is a critical variable that must be controlled in order to target the correct contaminant population for sampling and analysis.

In the groundwater sampling example discussed above, the difference between purged sampling and diffusion bag samples is their different sample supports in relation to the vertical stratification of adjacent populations. Inadvertently mixing two different populations through careless sample supports (when only one population is expected) creates misleading data. On the other hand, differing sample supports can produce non-comparable data sets, even if the samples are analyzed side-by-side by the exact same analytical method.

A number of newer analytical devices often used *in situ*, such as x-ray fluorescence (XRF), direct-push (DP) deployed laser-induced fluorescence (LIF), or DP-deployed membrane-interface probe (MIP) with specific detectors, have very small sample supports. Figure 5 illustrates trichloroethene (TCE) data generated by a MIP equipped with an electron capture detector (ECD) useful for chlorinated organics. Small sample supports can locate spatially discrete contaminant sources and migration conduits often missed by conventional monitoring wells. Monitoring wells have traditionally been placed "blind." Without a tool like the direct push MIP that develops the CSM (by detecting distinctly different populations) before well placement, data results and interpretation are highly uncertain. A well placed in the location represented in Figure 5 could be screened in any one of many possible configurations of depth and screen length, as illustrated by wells A, B and C. The TCE concentration expected from well configuration A could be very different from data produced by other configurations placed in the same bore hole.

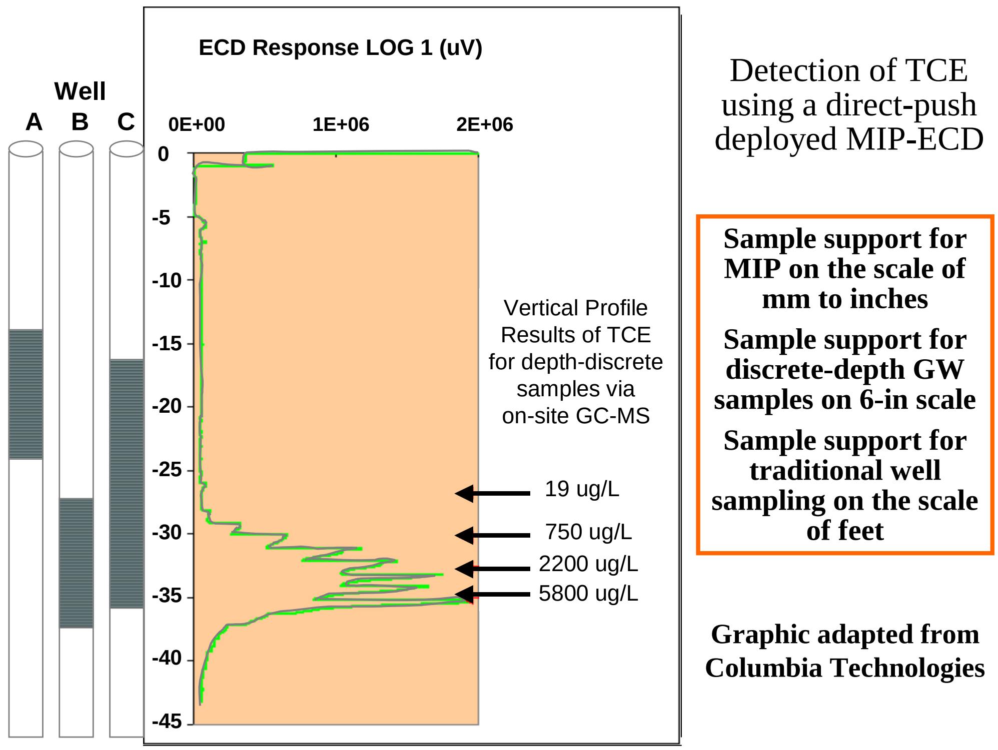

Figure 5. Different Sample Support Changes Analytical Results for GW

These different analyte concentrations are not the product of analytical uncertainty, but of sampling uncertainty. Data can be misleading if the sample support variable is not controlled.

Particle size is another physical aspect of sampling (i.e., sample support) that must be controlled when micro-scale heterogeneity is present (i.e., different populations are present in the same specimen). Table 1 summarizes a study that examined the relationship between the size of native soil particles and lead concentration at a firing range site (ITRC, 2003). The smaller the particle size, the higher the lead concentration. The bulk average concentration is about half the concentration of the smallest particles. Whether the bulk average is the correct sample support depends on the decision. Suppose the decision is to assess exposure risk from dust blowing off-site into local homes, sticking to children's fingers, which go into their mouths—the smallest particle size is representative of this exposure decision. Using the bulk average value as a default could underestimate true exposures by a factor of 2.

Table 1. Lead Concentration as a Function of Particle Size (after ITRC, 2003)

Soil Grain Size (Standard Sieve Mesh Size)	Soil Fraction-ization (%)	Pb Conc. in fraction by AA (mg/kg)	Lead Distribution
Greater than 3/8" (0.375")	18.85	10	0.20
Between 3/8 and 4-mesh"	4.53	50	0.24
Between 4- and 10-mesh	3.65	108	0.43
Between 10- and 50-mesh	11.25	165	2.00
Between 50- and 200-mesh	27.80	836	25.06
Less than 200-mesh	33.92	1,970	72.07
Totals	100%	927 (wt-averaged bulk)	100%

Particle size also impacts laboratory subsampling procedures. What particle sizes are preferentially captured by subsampling? A spoon-shaped scoop will retain a different mix of particle sizes than a narrow, flat spatula. Has the laboratory been advised what particle size they should target to maintain data representativeness for the specific decision(s) intended by the data users or project manager?

The phenomenon of highly concentrated particles encountered in Table 1 helps explain why smaller sample and **subsample volumes** produce more highly variable analytical results. A study in 1978 by the Department of Energy demonstrated this with soil from an area contaminated with americium-241 (Am-241, a radionuclide). A large volume of soil was sampled and containerized. It was carefully homogenized by drying, ball-milling, and sieving through a 10-mesh screen. Twenty subsamples each of various masses were taken and analyzed separately. The results are summarized in Table 2. Obviously, the larger the subsample, the less variable the results, and the much more reliably any single subsample result estimated the true mean (1.92 ppm) for the original sample.

Table 2. Subsampling Variability (adapted from Doctor and Gilbert, 1978)

Subsample Mass (g)	Range of Results for 20 Individual Subsamples (ppm)
1	1.01 to 8.00
10	1.36 to 3.43
50	1.55 to 2.46
100	1.70 to 2.30

A decision error could occur if a data user got the result of 8 ppm from a 1-gram subsample, and then assumed that the result represented the true concentration for the entire jar of sample (an error of about 400%). The error would be further compounded if that 8 ppm result was extrapolated to represent the concentration of Am-241 for a large portion of the site. Even with homogenization (which is never perfect), the smaller the subsample, the less likely that its result represents the average concentration for the original jar of soil. This is a problem for analytical chemistry: as instrumentation becomes more and more sophisticated, the mass of sample used by the laboratory to actually generate the analytical result is trending lower and lower. One gram is a standard sample size for soil digested for metals analysis. Results are viewed as "gold-plated" simply because of the accuracy of the determinative method (refer to Figure 4). But that is simply the last link in a chain of events made of weak links that are largely uncontrolled by standard practices for project planning and laboratory analysis.

SUMMARY

"Data representativeness" is the idea that we expect to extrapolate the data results from tiny matrix samples to draw conclusions about much larger populations. The tremendous mismatch in spatial scale between the volume of analytical samples and the volume of heterogeneous matrix targeted by projects decisions poses a severe challenge for "representative data." Data will be representative only if each sampling and analytical variable has been planned in advance to match the intended decision. CSMs are the tools used to develop and test hypotheses about what different contaminant populations may be present at the site, what variability is expected within each population, and how each population relates to the analytical and decision-making processes.

The U.S. EPA has articulated the Triad approach as a practical framework that synthesizes new technologies and advancing science with evolving regulatory and engineering practices governing site cleanup. The Triad approach rests on the foundation of managing decision uncertainty. Managing data uncertainty, especially that contributed by sampling uncertainty, is critical when decisions are being based on data. The second-generation framework offered by the Triad approach not only increases decision confidence, but also decreases project lifecycle costs by evolving the site conceptual model in real-time (using dynamic work strategies) to the extent feasible. The payoff for real-time decision uncertainty management is that Triad projects typically show lifecycle cost savings in the neighborhood of 30-50% as compared to first-generation strategies for site characterization, remediation, and monitoring.

REFERENCES

Doctor, P.G. and R.O. Gilbert. 1978. "Two Studies in Variability for Soil Concentrations: with Aliquot Size and with Distance." in "Selected Environmental Plutonium Research Reports of the Nevada Applied Ecology Group" (NAEG), Las Vegas, NV. pp. 405-423 and 439-442.

Huffman, R.L. 2002. Comparison of Passive Diffusion Bag Samplers and Submersible Pump Sampling Methods for Monitoring Volatile Organic Compounds in Ground Water at Area 6, Naval Air Station Whidbey Island, Washington. U.S. Geological Survey Water-Resources Investigations Report 02-4203. Available on-line at http://water.usgs.gov/pubs/wri/wri024203/

Interstate Technology and Regulatory Counsel (ITRC). 2003. Characterization and Remediation of Soils at Closed Small Arms Firing Ranges. Available: http://www.itrcweb.org/SMART-1.pdf

U.S. Department of Energy (DOE). 2001. Innovative Technology Summary Report: Adaptive Sampling and Analysis Programs (ASAPs). DOE/EM-0592. August. Available: http://cluin.org/download/char/ASAPs_ITSR_DOE-EM-0592.pdf

U.S. Environmental Protection Agency (EPA). 1993. Data Quality Objectives Process for Superfund: Interim Final Guidance. EPA540-R-93-071. Sept.

Vroblesky, D.A., and Peters, B.C., 2000, Diffusion sampler testing at Naval Air Station North Island, San Diego County, California, November 1999 to January 2000: U.S. Geological Survey Water-Resources Investigations Report 00-4812.