Advanced Design Application & Data Analysis for Field-Portable XRF

A Series of Web-based Seminars Sponsored by Superfund's Technology & Field Services Division

Session 2

Q&A for Session 1

Module 3.1 – Representativeness Part 1

3.1-1

How To . . .

- Ask questions
  - » "?" button on CLU-IN page
- Control slides as presentation proceeds
  - » manually advance slides
- Review archived sessions
  - » http://www.clu-in.org/live/archive.cfm
- Contact instructors

3.1-2

Notes

- When you registered, you were directed to this seminar's specific URL, which is the front page of today's seminar. The Front Page of the web cast contains a short abstract of today's session. We have also included pictures and short biosketches of the presenters. Please note the presenters' email addresses are hotlinked on that page in case you have any questions for one of them after today's presentation.

- For those of you joining us via the phone lines, we request that you put your phone on mute for the seminar. We will have Q&A sessions at which point you are welcome to take your phone off mute and ask the question. If you do not have a mute button on your phone, we ask that you take a moment RIGHT NOW to hit *6 to place your phone on MUTE. When we get to the question and answer periods you can hit #6 to unmute the phone. This will greatly reduce the background noises that can disrupt the quality of the audio transmission.

- Also, please do not put us on HOLD. Many organizations have hold music or advertisements that can be very disruptive to the call. Again, keep us on MUTE. DO NOT put us on HOLD.

- Also, if you experience technical difficulties with the audio stream, you may use the ? icon to alert us to the technical difficulties you are encountering. Please include a telephone number where you can be reached and we will try to help you troubleshoot your problem.

- Instructor contact information:

Deana Crumbling, U.S. EPA
Phone: (703) 603-0643
Fax: (703) 603-9135
Email: crumbling.deana@epa.gov

Robert Johnson, Argonne National Laboratory
Phone: (630) 252-7004
Fax: (630) 252-3611
Email: rlj@anl.gov

Stephen Dyment, U.S. EPA
Phone: (703) 603-9903
Fax: (703) 603-9135
Email: dyment.stephen@epa.gov

Q&A For Session 1 – Introduction and Basic XRF Concepts

3.1-3

Module 3.1:

Representativeness – Part 1

3.1-4

Data Collection Tied to Specific Decision

- Contaminant above background levels? (SI)
- Human health or ecological risks unacceptable? (RI)
- Contaminant concentrations above the cleanup criteria? If so, what should be done? (FS/RD)
- Should soil/sediment removal/treatment continue? be modified? or stop? (RA)
- Many kinds of data used to guide decisions.

3.1-5

Notes

- **Contaminant above background levels? (SI):** The site inspection (SI) phase of site characterization focuses on establishing whether or not contaminant levels are significantly above background so that they can be scored using the Hazard Ranking System model.

- **Human health or ecological risks unacceptable? (RI):** During the remediation investigation (RI), environmental data is collected to determine if the nature and extent of contamination present an unacceptable risk to human health or the environment. Environmental data is also collected during the RI to evaluate remediation technologies.

- **Contaminant concentrations above the cleanup criteria? If so, what should be done? (FS/RD):** During the feasibility study (FS) and remedial design (RD), environmental data are used to determine if contaminant concentrations are above preliminary remediation goals in the FS report or cleanup levels in the Record of Decision. Environmental data collected during RD are typically used to better estimate the volume of contaminated media exceeding contaminant concentrations.

- **Should soil/sediment removal/treatment continue? be modified? or stop? (RA):** Environmental data collected during the RA are used to determine if the cleanup objectives have been met and to monitor and optimize the operations of ongoing treatment systems.

Nirvana

- Representative, fast, cheap method able to run lots of samples and provide "definitive data"
- Reality Bites:
  - » Expensive, time consuming analytics
  - » Few samples collected
  - » Measurement errors & interpretation issues exist

3.1-6

Notes

- **Representative, fast, cheap method able to run lots of samples and provide "definitive data":** In a perfect world, the collection of representative data would be fast and inexpensive using methods that generate a large volume of data that is considered to be "definitive." What is the fundamental issue? If we had a cheap way of providing high density data quickly that was definitive relative to whatever decision that needed to be made, we would not be having this workshop.

- **Reality bites:** In reality we are asked to make decisions about a site from a very limited number of samples, whose results, at times, are subject to interpretation and error themselves. Traditionally, data collection is very expensive and uses time consuming analytical techniques. Because of the time and expense, only a few samples are collected. Relying on just a few samples leads to interpretation issues because of predictable measurement errors.

"Representative Sample" is Defined in Regulation

RCRA: "Representative sample means a sample of a universe or whole (e.g., waste pile, lagoon, ground water) which can be expected to exhibit the average properties of the universe or the whole (40 CFR 260.10)."

Evades several questions, such as…
…how do you take such a sample?

3.1-7

Notes

- **RCRA definition:** The Resource Conservation and Recovery Act regulations define the term representative sample as, "Representative sample means a sample of a universe or whole (e.g., waste pile, lagoon, ground water) which can be expected to exhibit the average properties of the universe or the whole (40 CFR 260.10)." The regulation makes it sound easy. But the regulation completely overlooks the physical reality that complexities (heterogeneity) exists on several spatial scales within a "waste pile, a lagoon, ground water". The difficulty lies in taking such a sample that represents the average of the universe or whole.

“Representative Sample” per RCRA

- Language unclear whether statistical or single physical sample is intended
  - » Common usage of “sample” implies a single grab sample can represent the whole
  - » Can we expect a single tiny sample to represent the average for a “whole”?
  - » Reasonable ONLY if homogeneity throughout the whole can be assumed

3.1-8

Notes

- **Language unclear whether statistical or single physical sample is intended:** The concept of “representativeness” is very vague and ill-defined for many working in the environmental field. This is because standard or regulatory definitions for “representativeness” also tend to be rather vague.

  » The RCRA solid waste regulations at 40 CFR §260.10 define a **representative sample** as: “a sample of a universe or whole (e.g., waste pile, lagoon, ground water) which can be expected to exhibit the average properties of the universe or whole.“ http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&sid=817591009b20d11a9dd16ef3f173c6aa&rgn=div8&view=text&node=40:23.0.1.1.1.2.1.1&idno=40

  » ASTM (consensus standard D 6044-96) defines a representative sample as “a sample collected in such a manner that it reflects one or more characteristics of interest (as defined by the project objectives) of a population from which it was collected.”

  It is not clear from these definitions whether the term “sample” refers to a statistical sample (made up of a number of individual specimens) or to a single sample, or whether the authors intended to allow either interpretation. A critical issue with the RCRA regulatory definition is that representativeness is defined in terms of an “average.” Operationalizing this definition for contaminated site cleanup poses problems. First, the extreme heterogeneity of environmental matrices and contaminants makes determination of a statistical “average” difficult and expensive. Second, how is it decided how to define the volume of the whole over which a property is to be averaged? Third, some environmental and engineering decisions (notably, those decisions involved with selecting and designing remedial systems) should not be made based on an "average," if that average encompasses wide variation.

In order to be useful for managing projects in the environmental field, the concept of representativeness must be made more concrete and meaningful. This can be done by simply adding the word "of." This adjusts the terminology and people's thinking to make it clear that data or other information must be representative of the intended decision or specific property under investigation. In this way, "representativeness" becomes linked to a concrete decision and decision unit rather than just an abstract "average." The ASTM definition seems to reflect this same kind of approach ("…reflects one or more characteristics of interest (as defined by the project objectives)…"

What is a "Representative Sample"?

- "A representative sample is one that answers a question about a population with a given confidence."
- "A sample that is representative for a specific question is most likely not representative for a different question."

From "A Methodology for Assessing Sample Representativeness" Charles Ramsey & Alan Hewitt in

Environmental Forensics, 6:71 – 75, 2005

ISSN: 1527 – 5922 print / 127 – 5930 online
DOI: 10.1080/15275920590913877

3.1-9

Notes

- **"A representative sample is one that answers a question about a population with a given confidence.":** A more useful definition of "representative sample" comes from literature. "A Methodology for Assessing Sample Representativeness," by Charles Ramsey and Alan Hewitt, focuses on the question being answered about the population from which the sample is taken.

- **"A sample that is representative for a specific question is most likely not representative for a different question.":** When this definition is linked to the question being answered then it also follows that if the question changes then a once "representative sample" may no longer be representative.

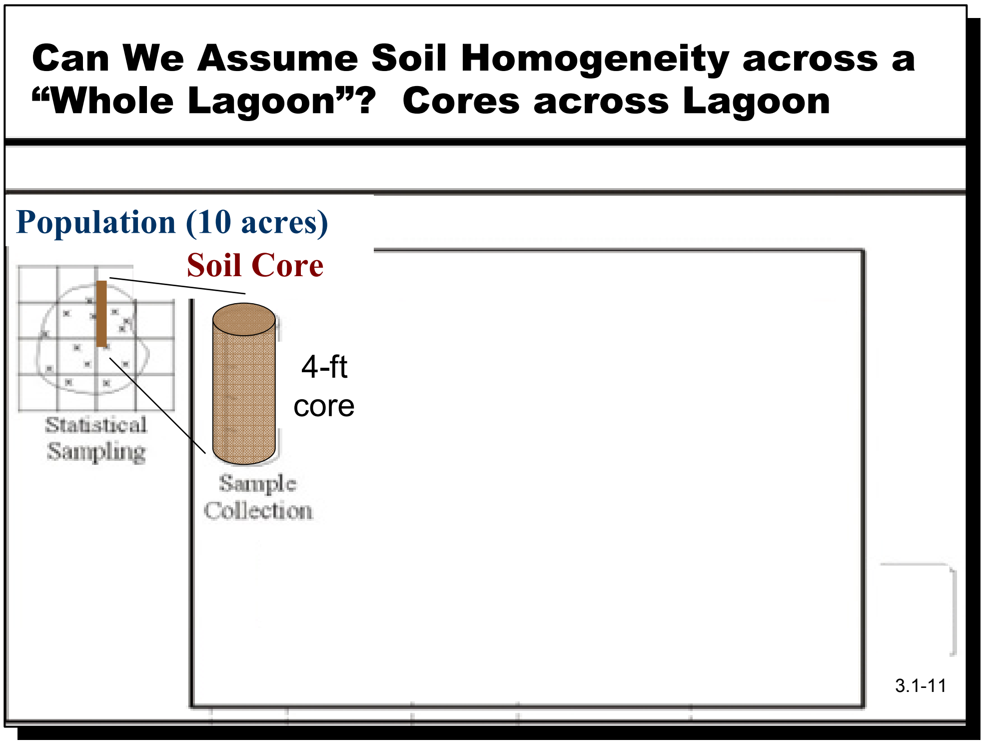
Representativeness: Answering a Question…Making a Decision

- Intended decision: Does the Lead (Pb) concentration in soil exceed a regulatory limit of 600 ppm for the top 4 ft of a 10-acre lagoon area?
- Samples taken as 4-ft cores. A 400-g soil sample is taken from the center of a core. A 1-g soil subsample is taken from the 400-g sample & analyzed for lead. The analytical result is 75 ppm Pb in soil.
- Does the Pb concentration over the top 4 ft of the 10-acre lagoon area exceed the 600 ppm limit?
- What does this analytical data result represent?

3.1-10

Notes

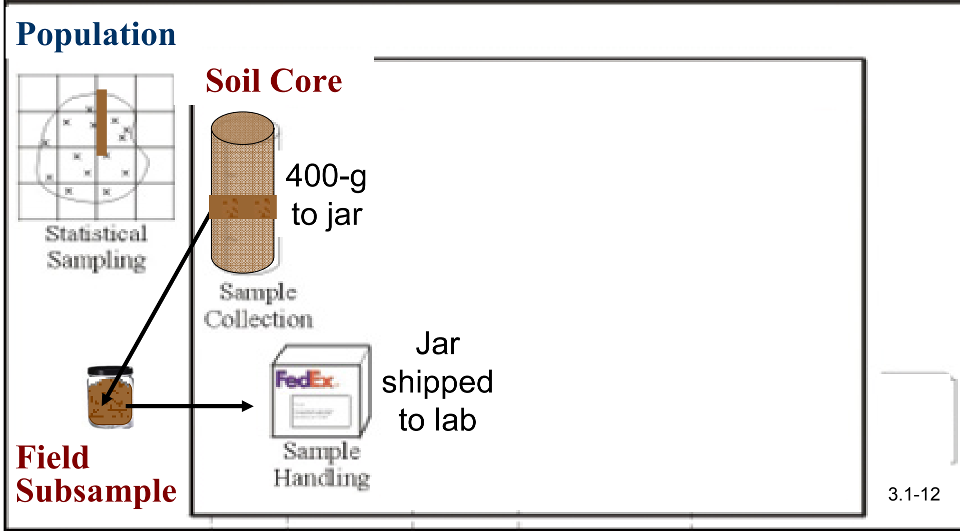
- **Data representativeness example:** The next several slides will explore the factors that complicate extrapolating the result of a 1-g subsample to a volume thousands of times larger.



Notes

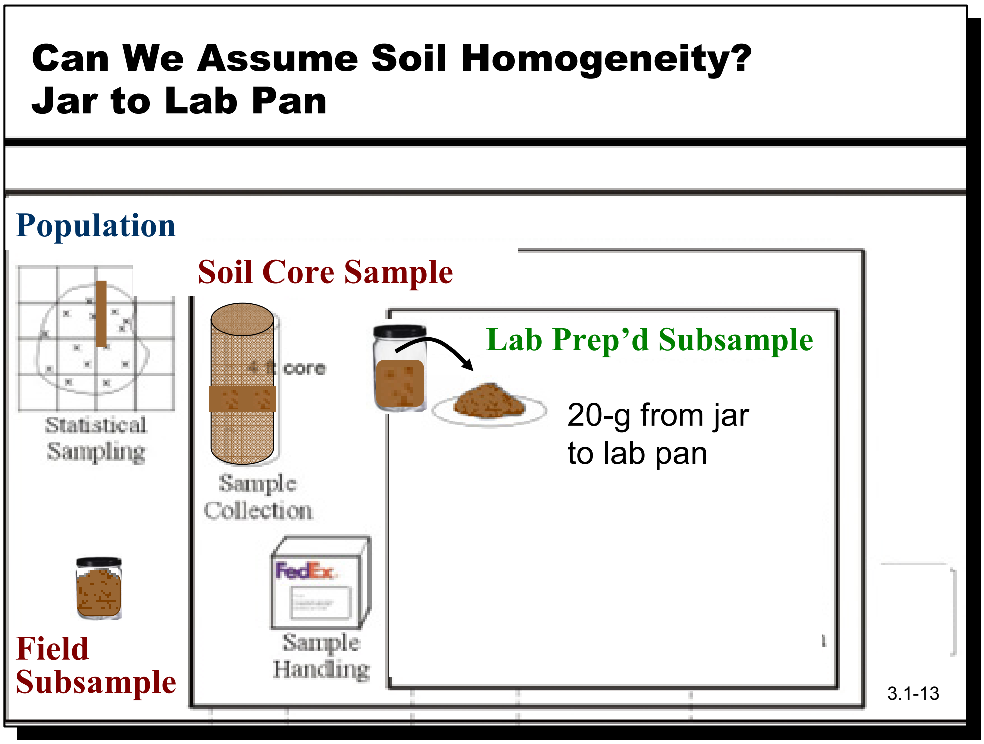
- Can a single sample be representative of average properties over 10 acres? Or must statistical sampling be used to gather a data set which can be averaged mathematically to produce measures that represent the overall average? Should we expect that the concentration at the top of a vertical core will be the same as the concentration at the bottom of a 4-ft core if the contaminant is released from the top?

Can We Assume Soil Homogeneity?
Core to Jar



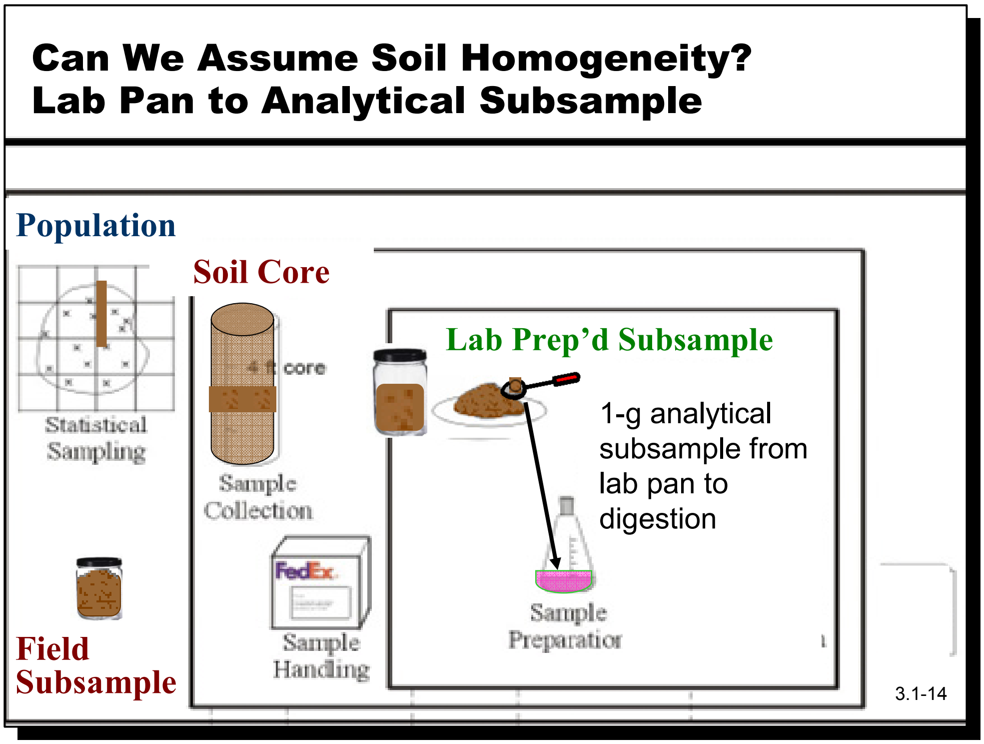
Notes

- Are we confident that a 400-g sample taken from near center of a core will be representative of the average over the whole core? Could it depend on the specific pollutant and how tightly it binds to soil particles?



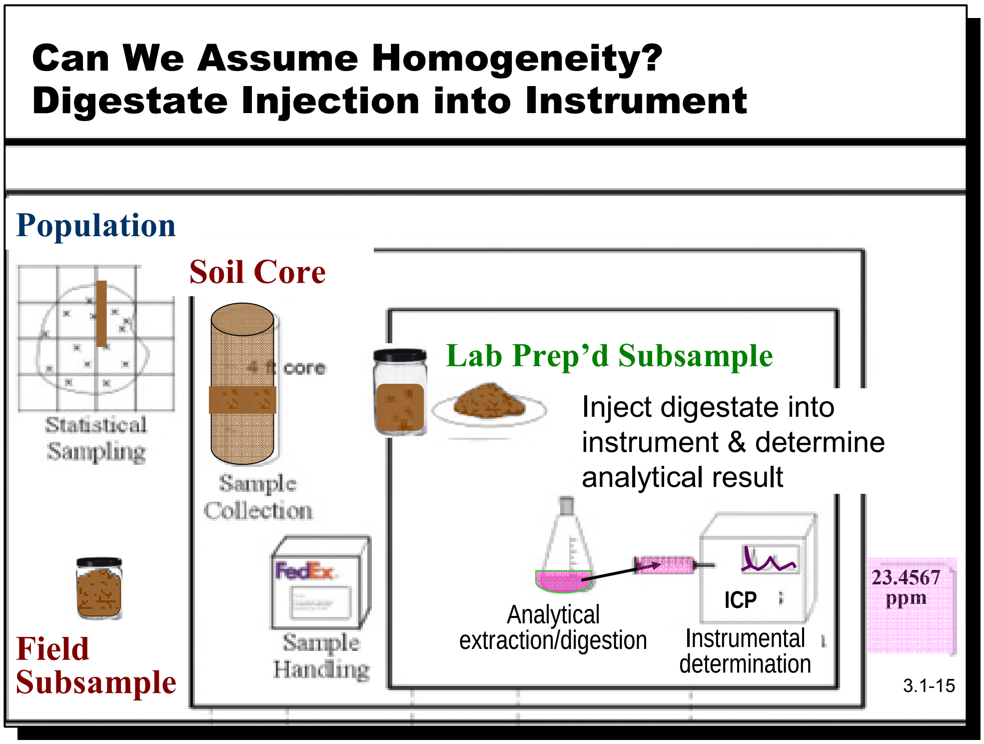
Notes

- Upon arrival in the lab, various options are possible. The subsample to be extracted or digested for analysis may be taken directly from the jar. Or all or part of the jar might be emptied into a lab pan. Then the sample may be stirred, or dried and sieved, or ground up, etc., prior to taking the analytical subsample.



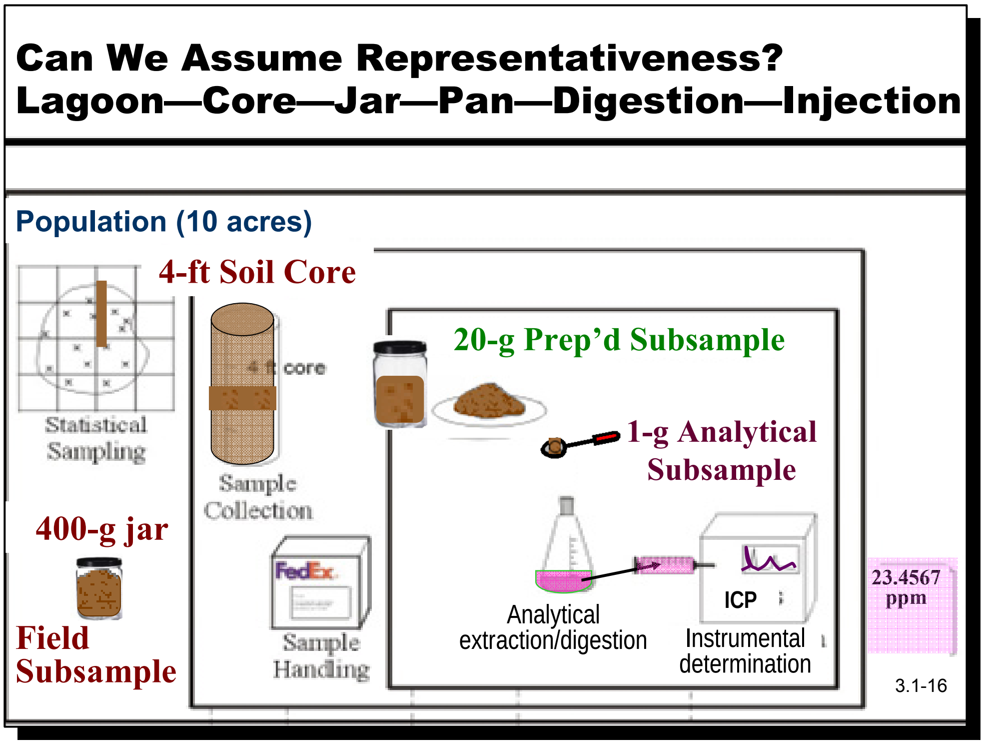
Notes

- The analytical subsample (the sample actually completely analyzed by the instrument) is usually quite small compared to the size of samples/subsamples further up the subsampling chain.



Notes

- Although not all of the solution generated during extraction or digestion is injected into the instrument for analysis, it is much easier to thoroughly mix/homogenize a solution than a solid material (like soil).



Notes

◆ Whether the analytical subsample is representative of the original sample taken in the field depends on the properties of the parent matrix, the sampling design (where and when samples were taken), and sampling/subsampling procedures. "Representativeness" here means that the analytical sample at the end of the subsampling chain has the same concentration as the average concentration over the large field volume considered to be the "whole." If an analytical subsample is representative, the result can be extrapolated to represent the average concentration of the whole.

Within-Sample Jar Variability: Micro-scale Sample Representativeness

Adapted from ITRC (2003)

Firing Range Soil Grain Size (Std Sieve Mesh Size)	Pb Concentration in fraction by AA (mg/kg)
Greater than 3/8” (0.375”)	**10**
Between 3/8” and 4-mesh	**50**
Between 4- and 10-mesh	**108**
Between 10- and 50-mesh	**165**
Between 50- and 200-mesh	**836**
Less than 200-mesh	**1,970**
Bulk Total	**927 (wt-averaged)**

The decision determines representativeness

3.1-17

Notes

- Lead concentration increases as soil particle size decreases. Which particle size is representative of the decision? The smallest particle sizes are representative of dust exposure decisions (like risk assessment characterization). On the other hand, the total soil profile is representative of decisions involving the soil as a bulk mass, such as TCLP testing for landfill decisions.

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

- Size conversions:

  » 3/8” = 0.375 in. = 9.525 mm

  » ASTM (US std) nominal aperture mesh size (mm):
    - 4-mesh = 4.76 mm
    - 10-mesh = 2 mm
    - 50-mesh = 0.297 mm = 0.3 mm
    - 200-mesh = 0.074 mm

Micro-scale Heterogeneity Causes Highly Variable Data Results for Replicates

Subsample Support (dried, ball-milled, sieved to <10-mesh)	Range of Results [for 20 individual subsamples (ppb)]	Coeff of Var. (CV)	Number of subsamples req'd to estimate true sample mean within a range of…	
			…± 25%* [ex: 1930 ± 25% = 1448 - 2412 ppb]	…± 10%* [ex: 1930 ± 10% = 1737 - 2123 ppb]
1 g	1010 - 8000	0.79	39	240

True mean for Am-241 in large sample known to be 1930 ppb

Adapted from DOE (1978)

* At 95% confidence

3.1-18

Notes

- A very large soil sample was analyzed using a radiological method for americium-241. The mean for the entire large sample was 1,930 ppb. The sample was then dried, milled and sieved to less than 10-mesh. Although this is not the most thorough sample preparation possible, it is far greater sample preparation than is generally performed by laboratories.
- After sample preparation, 20 1-g subsamples were taken from the large prepared sample. Each was analyzed, and the 20 sample results compiled. The range of results are provided in the 2nd column, and a statistical expression of variability among the 20 results, the coefficient of variation (CV), is provided in column 3. Columns 4 and 5 convey the impact of the level of variability as quantified by the CV. For 1-g subsamples, the CV was 0.79, which means that 39 replicate analyses would be required to reduce data uncertainty to +/- 25%, or 240 replicate analyses required to reduce data uncertainty to +/- 10%. In other words, if an estimate of the true mean of the sample were desired to be within +/- 10% of the true mean (1,930 ppb), to control for a CV = 0.79, 240 individual subsamples would have to be analyzed and averaged together to achieve a single result that has a 90% level of accuracy. Inaccuracy of +/- 10% means that the averaged mean will fall between 1,740 and 2,120 ppb, or +/- (1,930 x 0.10) ppb.
- Adapted from Source: 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.

  95% confidence for +/- % range.

The Smaller the Analytical Sample, the More Likely that the Result is Non-Representative

Subsample Support (dried, ball-milled, sieved to <10-mesh)	Range of Results [for 20 individual subsamples (ppb)]	Coeff of Var. (CV)	Number of subsamples req'd to estimate true sample mean within a range of… …± 25%* [ex: 1930 ± 25% = 1448 - 2412 ppb]	…± 10%* [ex: 1930 ± 10% = 1737 - 2123 ppb]
1 g	1010 - 8000	0.79	39	240
10 g	1360 - 3430	0.27	5	28
50 g	1550 - 2460	0.12	1	6
100 g	1700 - 2300	0.09	1	4

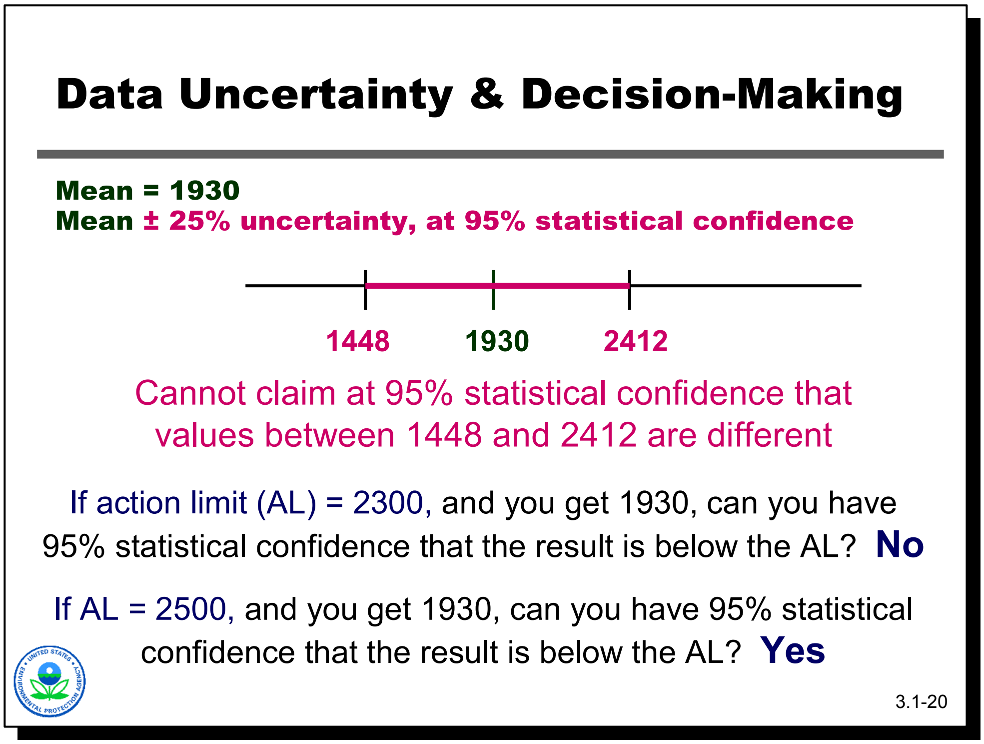
Adapted from DOE (1978)

How much confidence should be placed in any single result?

3.1-19

Notes

- In addition to the experimental trials using 1-g subsamples, additional rounds of sampling were done with 20 each of 10-g, 50-g, and 100-g subsample masses. The trend is obvious: the larger the subsample taken from a prepared sample, the lower the variability in replicate results, and the fewer replicate analyses would be needed to achieve a desired degree of data confidence.
- Advancing Technology Problem: Advancing analytical technologies and instrumentation are using smaller and smaller subsample aliquots. This phenomenon is reducing data quality by reducing the representativeness of individual sample results.
- Adapted from Source: 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. 95% confidence for +/- % range.



Notes

- When an acceptable uncertainty is set, such as +/- 25%, the meaning of that acceptable level of data uncertainty is that the measurement system (which includes sample collection and handling steps) is not expected to perform any better than +/- 25%. That means that each analytical result is bounded by an uncertainty interval of +25% on the upper side of the result, and an uncertainty interval of -25% on the lower side. Decision confidence at the specified degree of confidence is possible only if decision points fall outside the data point's uncertainty interval. Higher levels of data quality will have narrower uncertainty intervals.

- QC checks serve to verify that the measurement system is performing at the expected level of data quality. QC checks may also show that the measurement system is performing better than expected, which means that the widths of the uncertainty intervals are decreased.

The Relevance of QC to Decision-Making

Soil Result = 350 ppm
Lab Duplicate Result is allowed to be ± 30% RPD

This measurement system's acceptable "noise" means that values between 359 and 473 ppm cannot be distinguished as different

If project decisions are to be made at an action level (AL) = 400 ppm and the permissible system noise is +/- 30% RPD, should you expect to decide that a result of 350 ppm is below the AL?

No

3.1-21

Notes

- Note: RPD = relative percent difference. +/- X% RPD is a different measure than +/- X%. The equation used here for calculating RPD is the (1st value minus the 2nd value) divided by the average of the 2 values.
- 100 + 25% RPD = 129; 100 - 25% RPD = 78. 100 +/- 25% RPD = 78 to 129. 100 +/- 25% = 100 +/- 25 = 75 to 125.
- Whenever an acceptance limit for a QC check is set, such as +/- 30% RPD for lab duplicates, the meaning of that acceptance limit is that the measurement system (which includes sample handling steps) is not expected to perform any better than +/- 30% RPD at the level of the laboratory duplicates.
- If the precision of the measurement system is +/- 30% RPD for the laboratory duplicates, the expectation is that the sample result would have to be no more than 259 ppm before the sample can be considered to have a concentration less than the action level of 400 ppm. If it is desirable to show that a result of 350 ppm is less than an action level of 400 ppm, the measurement uncertainty must be reduced to less than 13% RPD. [((400-350)/375) x 100 = 13.3% RPD]
- Obviously, the "further" the sample result is from the action level, the less precise the measurement system needs to be to make confident decisions at the action level. If the result is 200 ppm, the width of the uncertainty interval can be up to about 65% RPD before the decision at 400 ppm becomes uncertain. [((400-200)/300) x 100 = 66.7% RPD]

Something to Think About: Lab Data Are Considered to be "Definitive"

Are lab data meeting all the requirements for definitive data?

DQOs for Superfund guidance

"For the data to be definitive, either analytical or total measurement error must be determined." (p. 43)

3.1-22

Notes

- Is the measurement or analytical error of data being reported to project managers in a way that allows them to take data uncertainty into account when evaluating data against decision criteria?
- Data Quality Objectives Process for Superfund: Interim Final Guidance (September 1993)

Measurement Error, Data Variability & Sample Representativeness

- We need a concept that captures the fact that the volume of the matrix is a determinant of measured concentrations and data variability (measurement error)
- That concept is called "**support**"
- The term and its definition appear in the *DQOs* for *Superfund* guidance and other EPA guidance documents for the waste cleanup programs

3.1-23

Notes

- **We need a concept that captures the fact that the volume of the matrix is a determinant of measured concentrations and data variability (measurement error):** Fortunately, such a concept does exist!
- **That concept is called "support":** Much of the remaining material in this module is dedicated to defining what "support" means.
- **The term and its definition appear in the *DQOs for Superfund* guidance and other EPA guidance for the waste cleanup programs:** The following guidance documents define "support:"
  - » Data Quality Objectives Process for Superfund: Interim Final Guidance (Sept. 1993)
  - » Guidance for Obtaining Representative Laboratory Analytical Subsamples from Particulate Laboratory Samples, EPA/600/R-03/027, November 2003
  - » Soil Sampling Quality Assurance User's Guide, 2nd Edition, EPA/600/8-89/046
  - » A Rationale for the Assessment of Errors in the Sampling of Soil, EPA/600/4-90/0-13

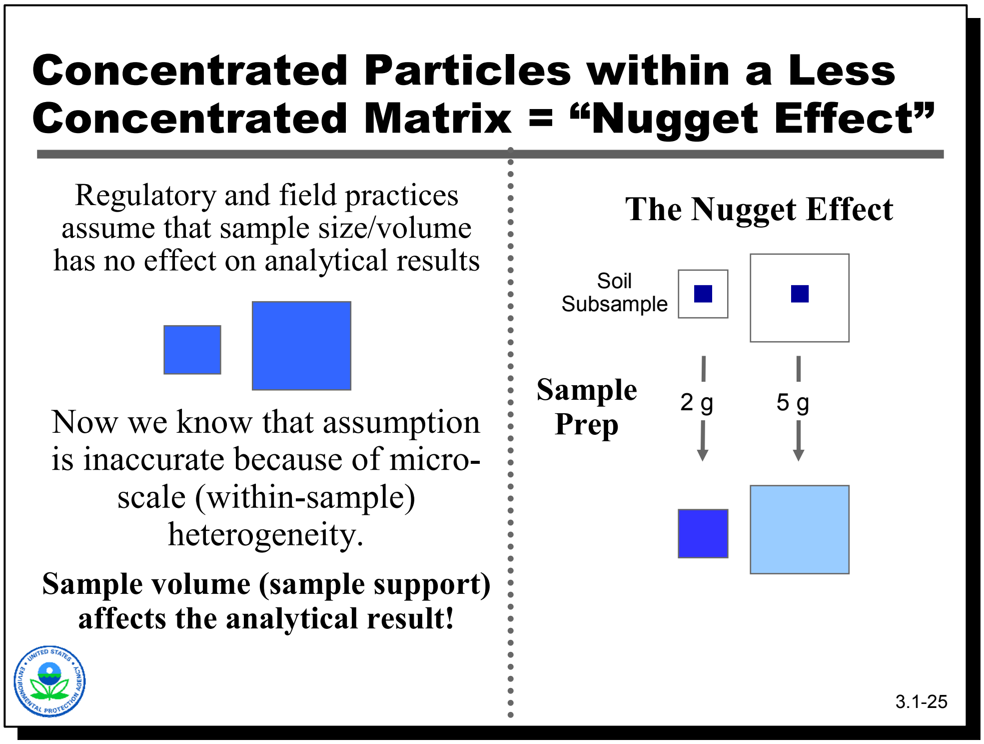
Sample Support

◆ **Sample Support** encompasses the physical properties of the sample that are relevant to the representativeness of the sample: the **size** (mass or volume), **shape** & **orientation** of a physical sample drawn from a matrix population (such as soil, sediment or water)

3.1-24

Notes

◆ **Sample support:** Sample support is the size (mass or volume), shape, and orientation of a physical sample drawn from a population (such as soil or water). The concept of "sample support" is necessary for the accurate collection and interpretation of data.



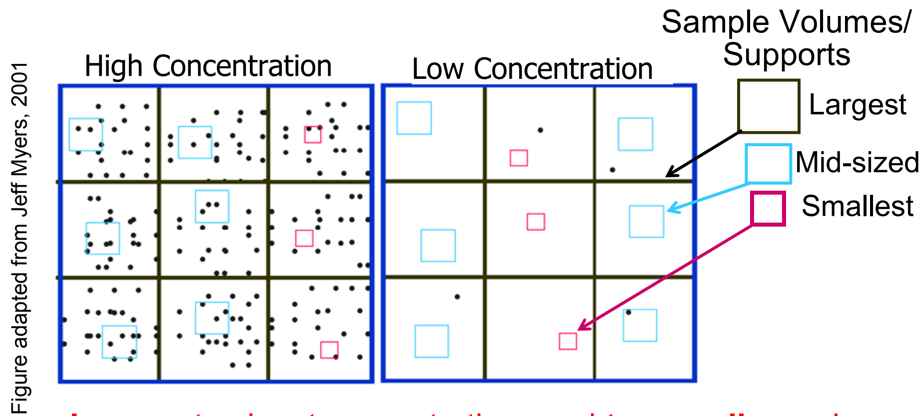
Notes

- The issue of "sample support" for heterogeneous environmental and waste matrices should make us reconsider the common (and usually unacknowledged) assumption that the reported concentration of an environmental sample should be the same no matter what volume of sample is collected.

- The volume of the sample is an important factor that influences the reported concentration for a sample, especially when contaminants are heterogeneously distributed throughout the parent matrix. All samples must be homogenized (through physical or chemical means) prior to analysis. For heterogeneous samples (that are affected by the nugget effect to a lesser or greater degree), the analytical result for a sample is determined by how much contaminant is captured in that sample, and how much cleaner matrix is contained in the sample (that serves to dilute the contaminant during homogenization). The nature of contaminant release to the environment (such as release to ground surface in the form of a powder or particulate) increases the probability of heterogeneity, as does contaminant solubility, mobility, and the age of the release. Obviously, environmental variables (such as precipitation, wind erosion, temperature, matrix composition) interact with the contaminants' properties to mitigate or aggravate heterogeneity. Contaminants that may at first have been more homogeneously released may become heterogeneously distributed throughout a matrix if their chemical properties cause them to preferentially partition onto mineral surfaces or into organic carbon that are themselves heterogeneously distributed, or inclusion of those matrix components into the analytical sample is variable or unpredictable.

- The issue of sample support is becoming an increasingly important determinant of analytical result as more sophisticated analytical technologies require smaller and smaller volumes of sample. At one extreme, sensor technologies currently under development will have miniscule sample supports, and data interpretation will be extremely difficult unless there is much greater awareness and management of sample support concepts.

Larger Sample Supports Produce More Consistent Data



Low contaminant concentrations and **too small** sample supports contribute to **lognormal populations!!**

3.1-26

Notes

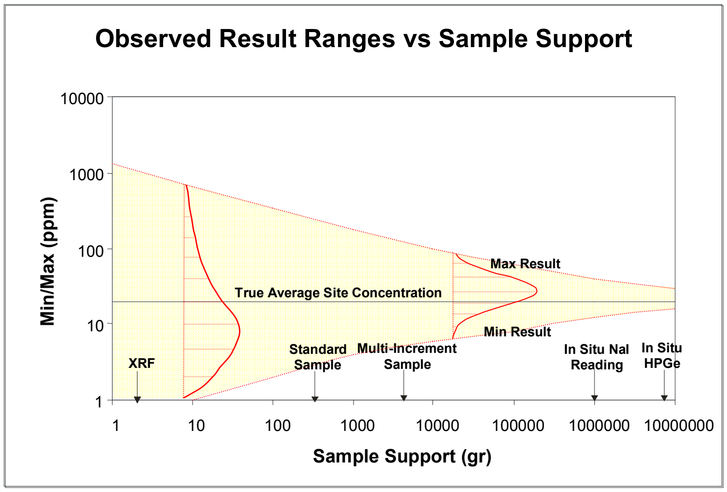
- The nugget effect can occur when contamination occurs in particulate form (such as explosives residues deposited as a powder or lead fragments in a firing range), or when contaminants partition onto mineral surfaces or organic carbon which are themselves heterogeneously distributed. Gy theory relates the size of the matrix particles to the sample support mass needed to be representative of the larger matrix volume. The volume of the sample is an important factor that influences the reported concentration for the sample, especially when contaminants are heterogeneously distributed throughout the parent matrix.

- Three different color-coded sample supports are illustrated in this figure. From largest to smallest sample support, the colors are black, light blue, and red. The dark dots (“particles”) represent higher contaminated small particles in a matrix of “cleaner” particles (not shown in the figure). The figure depicts the variable capture rates of the “dirty” particles for higher and lower contaminant concentrations and for different sample supports (volumes).

- Since smaller sample supports have a lower capture rate of contaminated particles, there is a higher rate of non-detects. On the other hand, when a contaminated particle is captured, the low volume of cleaner matrix causes the concentration to be higher after sample preparation. These two factors, a high number of non-detects/very low concentration samples, and very high concentrations in a few samples, produce the statistical lognormal frequency distribution common to environmental sampling.

- Specifying a regulatory threshold without specifying the sample support over which it applies (or at least recognizing that differences in sample support introduce variability into analytical data results) easily leads to widely different analytical results from one sample to the next. Since sample support is generally ignored in regulation, it is ignored in practice and the sample support is left to chance. This leads to uncontrolled (and usually undocumented) variations in sampling conditions and often widely varying results that are difficult to interpret. Unless the laboratory was in charge of field sampling and was involved in project planning and SAP/QAPP preparation, the laboratory cannot be held accountable for such variable results. The analytical result is probably correct from the standpoint of generating an accurate result on the analytical sample. Project planning was faulty for not ensuring that sample collection and handling procedures would produce samples representative of the decision.

Another Way to View Data Uncertainty as a Function of Sample Support



3.1-27

Notes

- This graph summarizes a study that looked at the range of results generated for different sample supports. This study is similar to the Am-241 study discussed earlier in this module. The analytical methods suitable for this analyte (uranium) have sample supports that range from 10,000 kg soil down to a couple of grams of soil. The larger the sample support, the more consistent multiple sample analyses are, and the closer each result is to the true mean. In contrast, very small sample supports create highly variable results. Many of the results are very far from the true mean.

- The superimposed frequency distributions illustrate how the lognormal distribution (the one on the left) is typical of small sample supports, whereas a more normal-shaped frequency distribution (the one on the right) is produced when larger sample supports are used. This directly connects to the previous two slides illustrating how smaller sample supports have a lower capture rate of contaminated particles, and a higher number of non-detect/very low values. The very high concentrations seen with small sample supports are due to the bias created when contaminated particles are captured within a small volume of cleaner matrix.

- It is important to know that the term “sample support” assumes that the sample volume is completely homogenized before any analytical subsampling is performed. If the sample volume is not appropriately homogenized prior to subsampling, then the volume of the analytical subsample is the actual sample support. On the graph, the sample support masses denoted as ”standard sample” and “ multi-increment sample” require that the full sample mass be completely homogenized before analytical sampling.

- The two analytical methods on the far right are in situ, walk-over detection technologies. The sample support reported for each is the volume of soil the detector/sensor is able to “see” all at once during a single analytical measurement.

- Abbreviations: XRF = X-ray fluorescence; NaI = sodium iodide detector; HPGe = high performance germanium detector.

Another Source of Data Variability: Short-scale, Between-Sample Heterogeneity

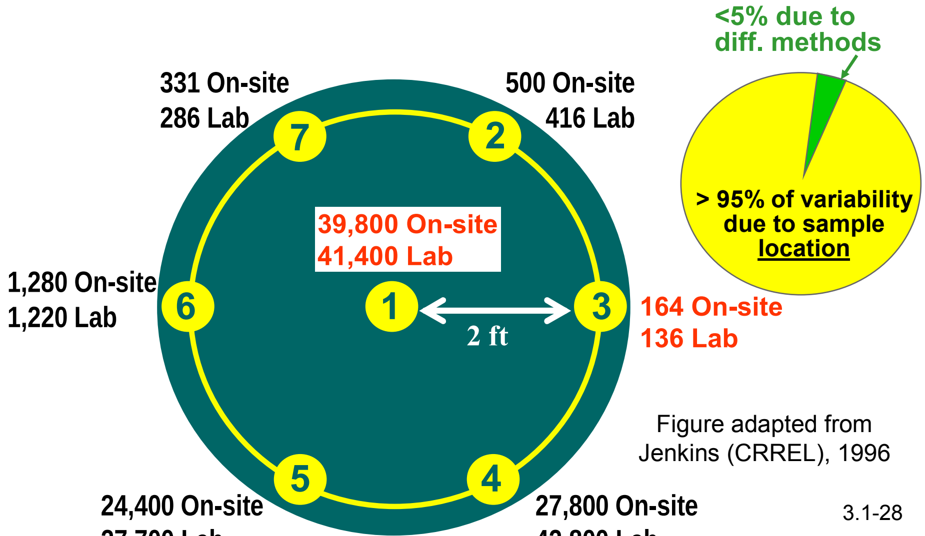


Figure adapted from Jenkins (CRREL), 1996

3.1-28

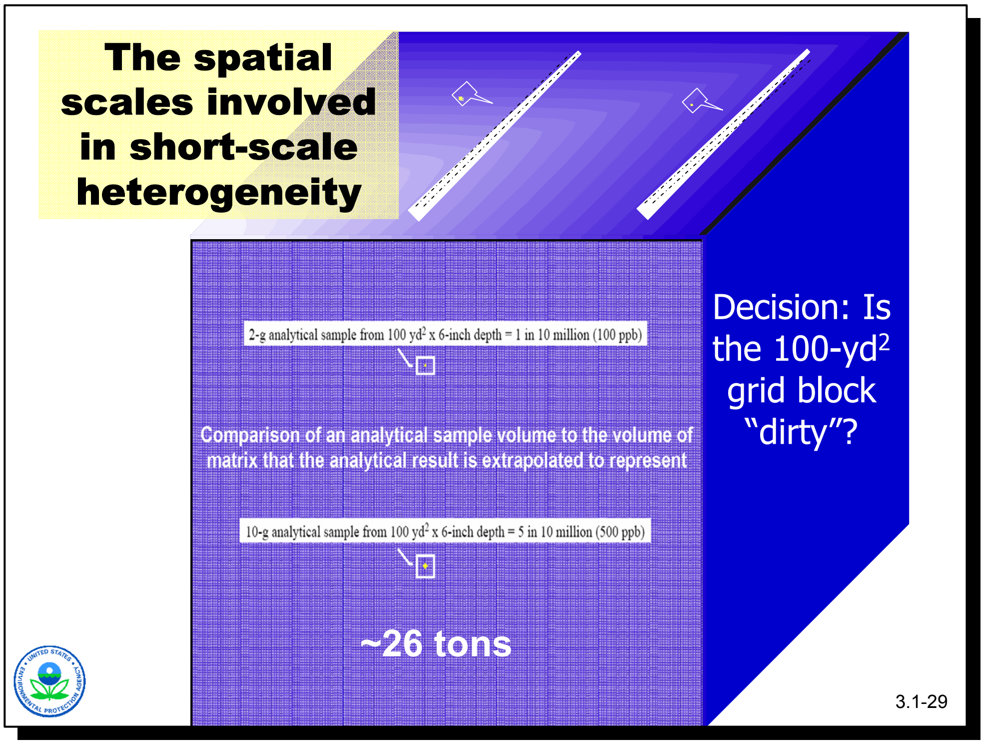
http://www.crrel.usace.army.mil/techpub/CRREL_Reports/reports/SR96_15.pdf

Notes

- This figure shows how spatial heterogeneity can be measured on a short spatial scale of several feet. The heterogeneity causes sample concentrations to differ by at least two orders of magnitude.

  This study examined the relative variability introduced in sample results by two different analytical techniques (in this case a standard lab analysis and a field analytical method) and by the short-scale spatial variability actually present in contamination concentrations (in this case explosive residues). The protocol selected seven samples from a 4-foot diameter circle. Those samples were split and analyzed two different ways. The larger graphic displays the observed results. Based on an analysis of variance, the smaller graphic apportions the observed variability in the complete data set (field analytics and lab data combined) between that contributed by differences in results produced by the two different techniques, and that contributed by spatial heterogeneity in this small area. The conclusion: the variability associated with spatial heterogeneity was 20 times greater than that associated with the analytical methods.

  Source: Example of characterizing sampling variability from USACE/CRREL work (Tom Jenkins) [see various reports at http://www.crrel.usace.army.mil/techpub/CRREL_Reports/html_files/Cat_X.html]. This example is from the Monite installation, which is contaminated with explosives residues (the facility reclaimed explosives from out-of-date munitions).



Notes

- The cube represents the volume of soil encompassed by 100 sq yds to a 6-inch depth (about 26 tons of soil). The higher dot represents the relative scale for a single 2-gram sample taken from that volume. The lower dot represents the relative scale for a single 10-gram sample.

  Assume a sample is designated as a "representative sample." The term is meaningless unless more information is supplied. What property is to be represented? The highest concentration of 1 contaminant, of all contaminants at once? The average contaminant concentration? What is the volume over which the result is to be represented?

  What is absolutely known is that the reported result is representative of the contaminant concentration of the analytical sample. Any extension of that concentration to other parts of matrix not analyzed must be supported by evidence that shows that extrapolation of the result to a larger volume is justified.

How to Pick the Right Sample Support

- **Sample Support** – the **size** (mass or volume), **shape** and **orientation** of the physical sample taken to ***represent*** a specific population of interest.
- Need to know what the population of interest is
- Need another concept: Decision Unit

3.1-30

Notes

- **Sample support:** The definition of sample support includes the population of interest.
- **Need to know what the population of interest is:** In order to select the appropriate sample support, the population of interest must be defined.
- **Need another concept:** The population of interest is defined by the decision to be made and includes the spatial area over which or the objects to which the decision will apply. This concept is called the decision unit.

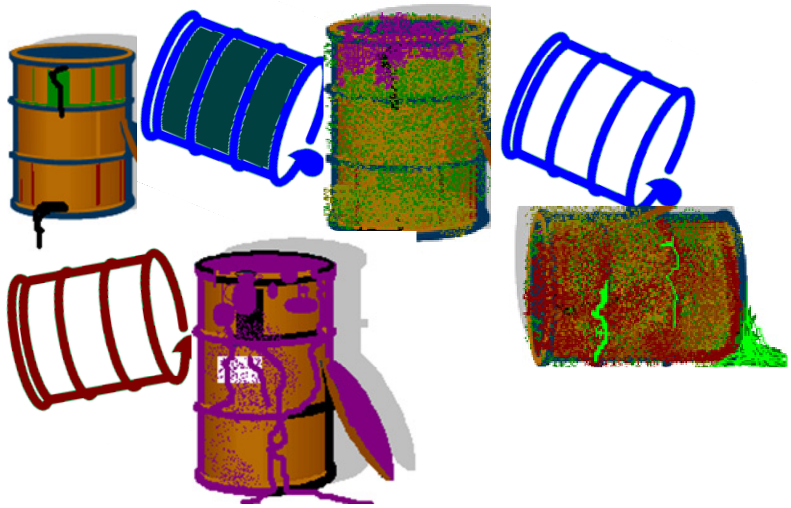
A Fundamental Concept for Ensuring Representativeness

- **Decision Unit:** An area, a volume, or a set of objects (e.g., ¼-acre area, bin of soil, set of drums) that is treated as a single unit when making decisions
- The decision unit may be a single item (such as a volume of soil)
- Or a decision unit may be a group of items united by a common property
- Examples: exposure unit, survey unit, remediation unit…

3.1-31

Notes

- **Decision unit:** A decision unit is an area or set of objects for which a decision needs to be made. The decision could be whether contamination is present in the unit or not, and if it is, whether it is at levels that exceed cleanup requirements.
- **The decision unit may be a single item (such as a volume of soil):** Examples of single decision units include a quarter acre or a city block.
- **Or a decision unit may be a group of items united by a common property:** Examples of a group of items is a set of storage drums or an abandoned laboratory with many small containers.
- **Examples:** Exposure units, survey units, remediation units, etc., are all examples of types of decision units. Decision units can have a temporal component as well as a spatial dimension.

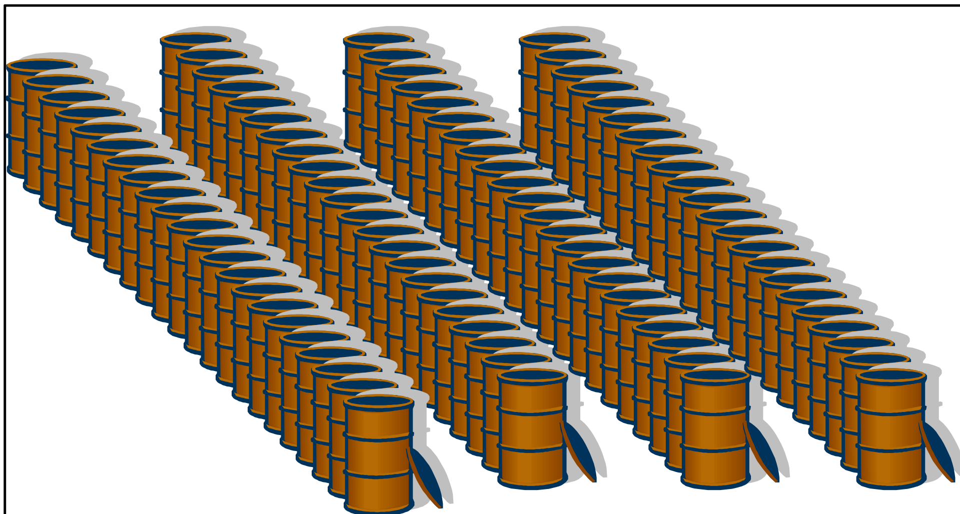


Valley of the Drums: These need to be characterized, transported, and disposed properly.
What is the decision unit? How do you sample it?

3.1-32

Notes

- Each drum is (or could be) completely different from the others in ways that can directly determine how the drum is handled and disposed.



Batch #1 Batch #2 Batch #3 Batch #4

40 drums were cleaned in batches of 20. You need to ensure the cleaning process worked.
What is the decision unit and how would you sample it?

3.1-33

Notes

- There is an expectation that each drum in a batch is the same as all other drums in the same batch, since they went through the same handling. Therefore, random statistical sampling of the batch is sufficient.

What is the Relationship Between Decision Unit and "Population"?

- **Population:** Set of objects or a volume of material sharing a common characteristic; can be synonymous w/ decision unit.
- Examples where they are not synonymous:
  - » 2 populations exist w/in a single decision unit:
    - – "clean" & "dirty" soil areas within a residential yard.
  - » A population is large enough so that more than 1 exposure unit is needed to cover it
    - – 50 acres are suspected to be clean; but the decision unit (exposure unit) is 1 acre. So there are 50 decision units in that "suspected clean" population.

3.1-34

Notes

- **Population:** A **population** is a set of objects or an area that shares some common characteristic. The way we will use population is often synonymous with decision units, however it need not be. A population might, in fact, be divided into several distinct decision units.

- A **population distribution** refers to the distribution (think histogram) of some population parameter that is of particular interest (e.g., contamination concentration). For example, the population might be everyone who lives in the Chicago metro region, the parameter of interest could be their height, and the population distribution for height would refer to the distribution of height present in Chicago's population.

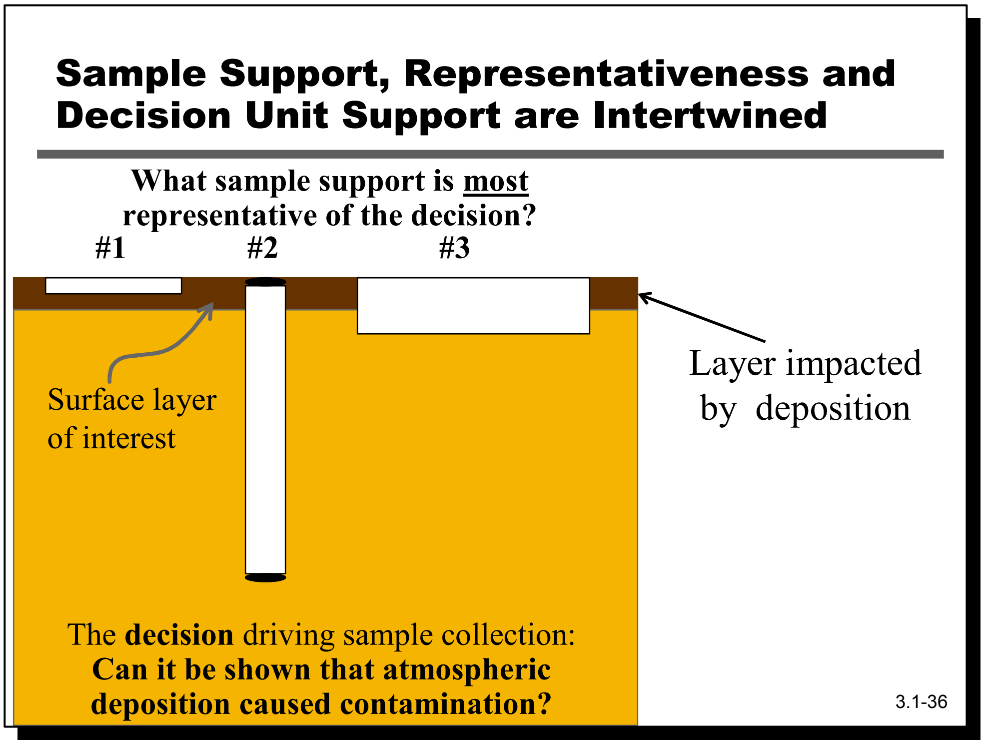
Sampling Design Requires A Progression of Supports

- Decision Unit Support – the spatial dimensions, mass, particle size or other physical properties that characterize the population of interest
  - » the decision DEFINES the population of interest, and
  - » the population DEFINES the properties of samples
- The sample support must mimic (on a small spatial scale) the decision unit support (on the larger spatial scale of the decision)

3.1-35

Notes

- **Sampling design requires a progression of supports:** Sampling program design starts “high” (at the intended decision) and works downward to representative sample support and analytics.



Notes

- This series of slides will illustrate concepts related to "sample support." These concepts are presented in a simplified form and do not attempt to portray the more exacting aspects of this topic.

- The panel on the left illustrates how sample volume and orientation must be selected to be representative of the decision to be made. Any of the 3 samples might be argued to represent true site conditions, but only one can be argued to be representative of site conditions in the context of the decision (atmospheric deposition).

  Color Key for left panel:

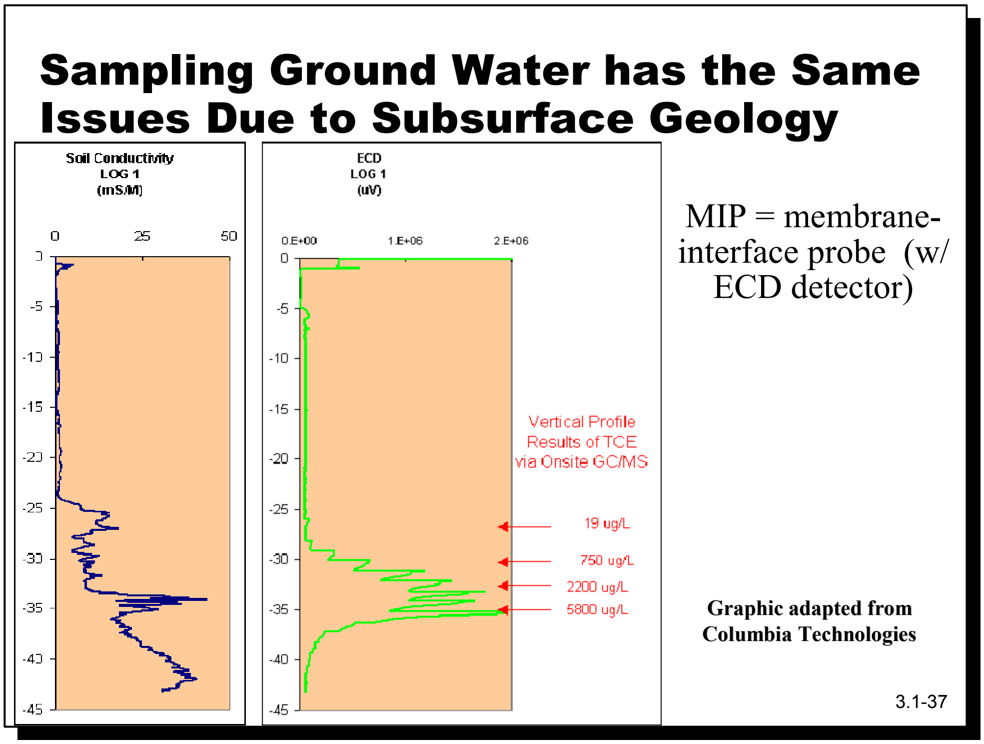
  » Dark brown depicts surface soil impacted by surface deposition of lead from the atmosphere.

  » Light brown depicts soil that would not be expected to be impacted by this atmospheric deposition.

  » White areas depict the volume and orientation of material removed that becomes the "sample."

  Keep in mind that the entire sample is homogenized prior to subsampling for analysis.

The sample support (the physical dimensions of the sample) for Sample #1 would be representative of the matrix impacted by atmospheric deposition, but the sample supports of samples #2 and #3 would not be. Sample support #3 illustrates the importance of strict control over sample support in scenarios where careful stratification of populations is required to avoid biasing results by including non-representative sample. Even though the general orientation of sample collection in #3 is similar to #1, the concentration of lead in sample #3 would be expected to be “diluted” by the inclusion of “cleaner” soil from a non-representative layer into the sample.



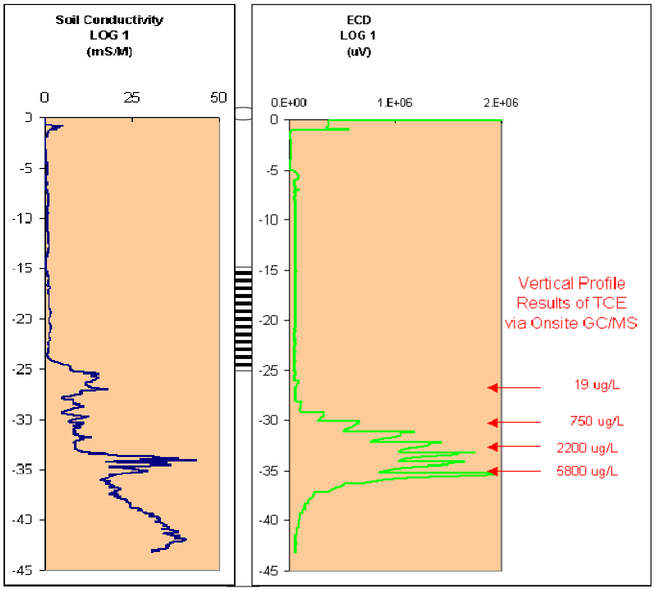
Notes

- Direct-push MIP-ECD taking readings every 2 inches going down to create a vertical profile of contamination in the subsurface.

  Soil conductivity results suggest transitions from sandy matrix to clay matrix (higher conductivity in clayey soil). The 7- to 8-ft wide band of contamination is associated with a clay layer in the subsurface. Small, discrete ground water (GW) samples (i.e., very small sample support) representative of point concentrations were collected using the DP probe and analyzed using GC-MS.

  What analytical results (low, medium, or high laboratory results) would be expected if a monitoring well were screened over the various intervals depicted in the slide? (Keep in mind that clay layers may be rather non-permeable to water flow as compared to sandier layers.)

Ground Water Sample Support (#1)



MIP = membrane-interface probe (w/ ECD detector)

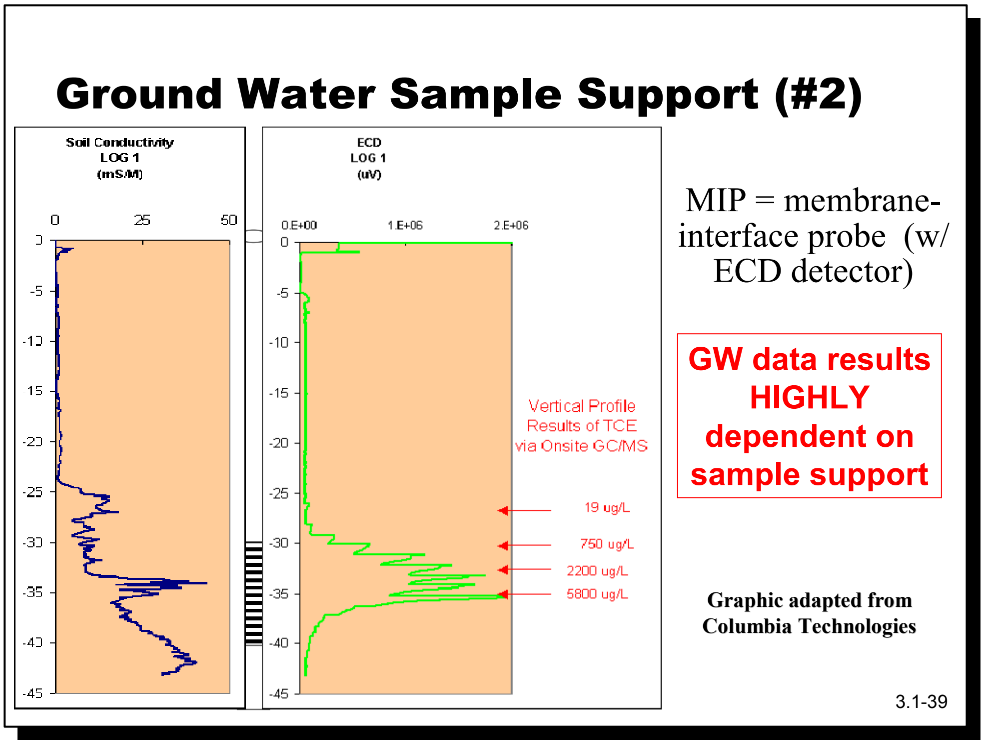
GW data results HIGHLY dependent on sample support

Graphic adapted from Columbia Technologies

3.1-38

Notes

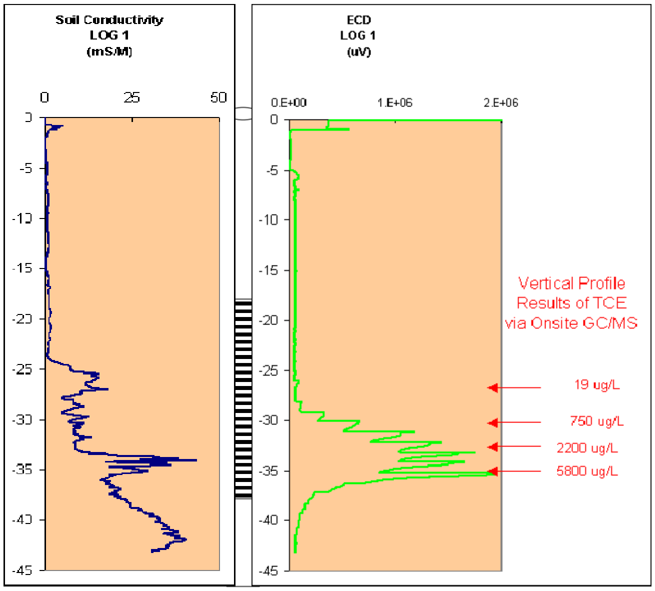
- What analytical results (low, medium, or high lab results) would be expected if a monitoring well were screened over the various intervals depicted in the graphic? (Keep in mind that clay layers may be rather non-permeable to water flow as compared to sandier layers.)



Notes

- What analytical results (low, medium, or high lab results) would be expected if a monitoring well were screened over the various intervals depicted in the graphic? (Keep in mind that clay layers may be rather non-permeable to water flow as compared to sandier layers.)

Ground Water Sample Support (#3)



MIP = membrane-interface probe (w/ ECD detector)

GW data results HIGHLY dependent on sample support

Graphic adapted from Columbia Technologies

3.1-40

Notes

- What analytical results (low, medium, or high lab results) would be expected if a monitoring well were screened over the various intervals depicted in the graphic? (Keep in mind that clay layers may be rather non-permeable to water flow as compared to sandier layers.)

The Biggest Cause of Misleading Data

3.1-41

Notes

- We may try to pretend that we can sample soil and ground water as if it were homogeneous…as if the size of decision units and sample support didn't matter. But Mother Nature isn't required to go along with our self-deception. No matter how accurate an analytical method may be, getting the right result on a non-representative sample will still give the wrong answer. And we will waste resources cleaning up matrix that doesn't need cleaning up, calling something clean when it actually needs remediation, designing remedial systems that don't work, and making erroneous decisions about risk.

Q&A – If Time Allows

3.1-42