
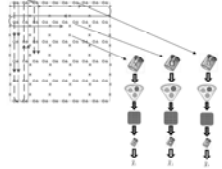


1

Welcome – Thanks for joining
this ITRC Training Class



**Soil Sampling and Decision Making Using
Incremental Sampling Methodology (ISM)
Part 1 – Principles, Systematic Planning, and
Statistical Design**



Web-Based Document at:
<http://www.itrcweb.org/ISM-1/>

**Incremental Sampling Methodology Technology Regulatory
and Guidance Document (ISM-1, February 2012)**

Sponsored by: Interstate Technology and Regulatory Council (www.itrcweb.org)
Hosted by: US EPA Clean Up Information Network (www.cluin.org)

When sampling soil at potentially contaminated sites, the goal is collecting representative samples which will lead to quality decisions. Unfortunately traditional soil sampling methods don't always provide the accurate, reproducible, and defensible data needed. Incremental Sampling Methodology (ISM) can help with this soil sampling challenge. ISM is a structured composite sampling and processing protocol that reduces data variability and provides a reasonable estimate of a chemical's mean concentration for the volume of soil being sampled. The three key components of ISM are systematic planning, field sample collection, and laboratory processing and analysis. The adequacy of ISM sample support (sample mass) reduces sampling and laboratory errors, and the ISM strategy improves the reliability and defensibility of sampling data by reducing data variability.

ISM provides representative samples of specific soil volumes defined as Decision Units. An ISM replicate sample is established by collecting numerous increments of soil (typically 30 to 100 increments) that are combined, processed, and subsampled according to specific protocols. ISM is increasingly being used for sampling soils at hazardous waste sites and on suspected contaminated lands. Proponents have found that the coverage afforded by collecting many increments, together with disciplined processing and subsampling of the combined increments, yields consistent and reproducible results that in most instances have been preferable to the results obtained by more traditional (e.g. discrete) sampling approaches.

This 2-part training course along with ITRC's web-based [Incremental Sampling Methodology Technical and Regulatory Guidance Document \(ISM-1, 2012\)](http://www.itrcweb.org/ISM-1/) is intended to assist regulators and practitioners with the understanding the fundamental concepts of soil/contaminant heterogeneity, representative sampling, sampling/laboratory error and how ISM addresses these concepts. Through this training course you should learn:

- basic principles to improve soil sampling results
- systematic planning steps important to ISM
- how to determine ISM Decision Units (DU)
- the answers to common questions about ISM sampling design and data analysis
- methods to collect and analyze ISM soil samples
- the impact of laboratory processing on soil samples
- how to evaluate ISM data and make decisions

In addition this ISM training and guidance provides insight on when and how to apply ISM at a contaminated site, and will aid in developing or reviewing project documents incorporating ISM (e.g., work plans, sampling plans, reports). You will also be provided with links to additional resources related to ISM.

The intended users of this guidance and training course are state and federal regulators, project managers, and consultant personnel responsible for and/or directly involved in developing, identifying or applying soil and sediment sampling approaches and establishing sampling objectives and methods. In addition, data end users and decision makers will gain insight to the use and impacts of ISM for soil sampling for potentially contaminated sites.

Recommended Reading: We encourage participants to review the ITRC ISM document (<http://www.itrcweb.org/ISM-1/>) prior to participating in the training classes. If your time is limited in reviewing the document in advance, we suggest you prioritize your time by reading the Executive Summary, Chapter 4 "Statistical Sampling Designs for ISM," and Chapter 7 "Making Decisions Using ISM Data" to maximize your learning experience during the upcoming training classes.


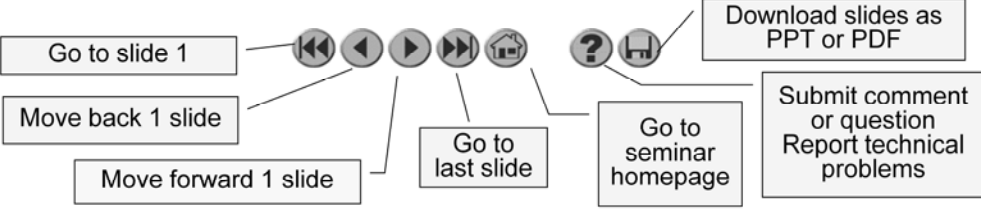
ITRC (Interstate Technology and Regulatory Council) www.itrcweb.org

Training Co-Sponsored by: US EPA Technology Innovation and Field Services Division (TIFSD) (www.clu-in.org)

ITRC Training Program: training@itrcweb.org; Phone: 402-201-2419

2

Housekeeping

- ▶ Course time is 2¼ hours
- ▶ Question & Answer breaks
 - Phone - unmute #6 to ask question out loud
 - Simulcast - ? icon at top to type in a question
- ▶ Turn off any pop-up blockers
- ▶ Move through slides
 - Arrow icons at top of screen
 - List of slides on left
- ▶ Feedback form available from last slide – **please** complete before leaving
- ▶ This event is being recorded

**Copyright 2015 Interstate Technology & Regulatory Council,
50 F Street, NW, Suite 350, Washington, DC 20001**


Although I'm sure that some of you are familiar with these rules from previous CLU-IN events, let's run through them quickly for our new participants.




We have started the seminar with all phone lines muted to prevent background noise. Please keep your phone lines muted during the seminar to minimize disruption and background noise. During the question and answer break, press *6 to unmute your lines to ask a question (note: *6 to mute again). Also, please do NOT put this call on hold as this may bring unwanted background music over the lines and interrupt the seminar.


You should note that throughout the seminar, we will ask for your feedback. You do not need to wait for Q&A breaks to ask questions or provide comments using the ? icon. To submit comments/questions and report technical problems, please use the ? icon at the top of your screen. You can move forward/backward in the slides by using the single arrow buttons (left moves back 1 slide, right moves advances 1 slide). The double arrowed buttons will take you to 1st and last slides respectively. You may also advance to any slide using the numbered links that appear on the left side of your screen. The button with a house icon will take you back to main seminar page which displays our presentation overview, instructor bios, links to the slides and additional resources. Lastly, the button with a computer disc can be used to download and save today's presentation slides.

ITRC (www.itrcweb.org) – Shaping the Future of Regulatory Acceptance



- ▶ Host organization 
- ▶ Network
 - State regulators
 - All 50 states, PR, DC
 - Federal partners




 - ITRC Industry Affiliates Program


 - Academia
 - Community stakeholders
- ▶ Disclaimer
 - Full version in “Notes” section
 - Partially funded by the U.S. government
 - ITRC nor US government warranty material
 - ITRC nor US government endorse specific products
 - ITRC materials copyrighted
- ▶ Available from www.itrcweb.org
 - Technical and regulatory guidance documents
 - Internet-based and classroom training schedule
 - More...

The Interstate Technology and Regulatory Council (ITRC) is a state-led coalition of regulators, industry experts, citizen stakeholders, academia and federal partners that work to achieve regulatory acceptance of environmental technologies and innovative approaches. ITRC consists of all 50 states (and Puerto Rico and the District of Columbia) that work to break down barriers and reduce compliance costs, making it easier to use new technologies and helping states maximize resources. ITRC brings together a diverse mix of environmental experts and stakeholders from both the public and private sectors to broaden and deepen technical knowledge and advance the regulatory acceptance of environmental technologies. Together, we’re building the environmental community’s ability to expedite quality decision making while protecting human health and the environment. With our network of organizations and individuals throughout the environmental community, ITRC is a unique catalyst for dialogue between regulators and the regulated community.

For a state to be a member of ITRC their environmental agency must designate a State Point of Contact. To find out who your State POC is check out the “contacts” section at www.itrcweb.org. Also, click on “membership” to learn how you can become a member of an ITRC Technical Team.

Disclaimer: This material was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof and no official endorsement should be inferred.

The information provided in documents, training curricula, and other print or electronic materials created by the Interstate Technology and Regulatory Council (“ITRC” and such materials are referred to as “ITRC Materials”) is intended as a general reference to help regulators and others develop a consistent approach to their evaluation, regulatory approval

Meet the ITRC Trainers



Annette Dietz
Portland State
University (formerly
Oregon DEQ)
Portland, OR
503-725-2328
acdietz@pdx.edu



Robin Boyd
AECOM Environment
Honolulu, HI
808-356-5376
robin.boyd
@aecom.com



Deana Crumbling
U.S. Environmental
Protection Agency
Washington, DC
703-603-0643
crumbling.deana
@epa.gov



Kelly Black
Neptune and Company, Inc.
Lakewood, CO
720-746-1803
kblack@neptuneinc.org

Annette C. Dietz, Ph.D., has been an instructor in the Department of Civil and Environmental Engineering at Portland State University in Portland, Oregon since 2015. From 2010 to 2015, she was the Cleanup Program Coordinator for the Oregon Department of Environmental Quality (OR DEQ) located in Portland, Oregon. She was a technical and policy expert for cleanup program activities, oversaw the maintenance and development of policy and guidance for cleanup project work, coordinated program functions performed by technical staff in regional offices, and provided outreach and training support for external stakeholders. Before joining OR DEQ in 2010, Annette worked as an environmental consultant managing and performing site investigation and remediation projects. While at OR DEQ, she was Oregon's Point of Contact (POC) to ITRC and a member of the ITRC Incremental Sampling Methodology team. Annette earned a bachelor's degree in Spanish and Global Studies in 1993, a master's in environmental engineering in 1996 and a doctoral degree in environmental engineering in 2000, all from the University of Iowa in Iowa City, Iowa.

Deana Crumbling is an Environmental Scientist in the Technology Innovation section of EPA's Office of Solid Waste and Emergency Response in Washington, DC. Since 1997 she has focused on the topics of dynamic work plans, field analysis and other analytical chemistry technologies, sampling designs and data uncertainty management. She has taught many classroom and webinar courses in those topics. Before coming to EPA, she worked for 2 years as a risk assessor for an environmental consulting firm, for 2 years as a lab and safety manager and adjunct faculty for a college, for 1 year as a science consultant for an environmental attorney, and 1 year in the Pennsylvania State Hazardous Site Cleanup Program. She worked for 20 years as a hospital-based clinical laboratory analyst and trainer. Deana was a team member in the ITRC Site Characterization and Monitoring Team from 2002 to 2004, and now a member of the Incremental Sampling Methodology Team since 2009. She earned a bachelor's in biochemistry and a bachelor's in psychology from Lebanon Valley College in Lebanon, Pennsylvania, in 1989, and a master's in environmental science from Drexel University in Philadelphia, Pennsylvania, in 1997.

Robin Boyd is Senior Project Manager with AECOM in Honolulu, Hawaii and serves as Technical Development Manager for the Honolulu office. Robin is a specialist in the use of incremental sampling for both surface and subsurface remedial investigations and remediation. Since 1989 Robin has gained experience with investigating hazardous waste sites including the development and implementation of innovative technologies for the remediation of soil and ground water. He has designed and implemented over 25 incremental sampling programs of various types and conducted numerous training classes on the use of incremental sampling. He has attended both Francis Pitard's and Chuck Ramsey's classes on incremental sampling and Gy's Sampling Theory (which is the cornerstone of ISM). He is an active member of the ITRC Incremental Sampling Methodology team. Robin earned both a bachelor's degree in 1979 and master's degree in 1981 in geology and geophysics from the University of Wisconsin in Madison. Robin also serves as a registered professional geologist in several states.

Kelly Black is a Senior Environmental Statistician with Neptune and Company located in Denver, Colorado. She has worked in the environmental arena for the past two decades and has broad experience with environmental problems from systematic planning, through development of sampling plans, to data analysis. Her work focuses on using creative sampling solutions to optimize decision making within the ever-present budgetary and practical constraints of environmental projects. Kelly has been teaching environmental statistics and quality assurance/quality control workshops for more than 15 years. She has been active in ITRC since 2004, first participating on the Brownfields team and then as co-lead for the statistics subgroup on the Incremental Sampling Methodology team. Kelly earned a bachelor's degree in statistics with minors in Economics and Sociology from Montana State University in Bozeman, Montana, in 1986, and a master's in statistics from Carnegie Mellon University in Pittsburgh, Pennsylvania in 1987.

Why Collect Soil Samples?

Representative Data:

- ▶ Accurate
- ▶ Reproducible
- ▶ Defensible



....but how do we get it?

Incremental Sampling Methodology (ISM)

.....may be your answer.....

Picture Reference: <http://www.swrcb.ca.gov/rwqcb2/brownfields.shtml>

Are Your Samples Representative?

- ▶ How fully do you plan your sampling event?
- ▶ Are you confident in your sample results?
- ▶ How representative are your samples?
- ▶ Do you understand the distribution?
- ▶ How reproducible are your data?



How are your decisions made when you review soil sampling data?

Sample Results - #%*&^%!

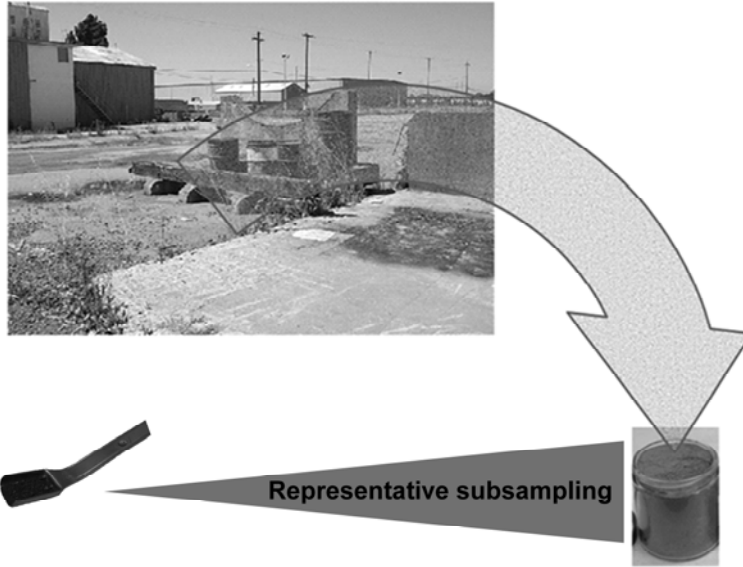


Discrete samples – if you find a hotspot you step out... but if you get a ND – are you done?

With a small number of discrete samples how well did you define the extent of contamination to begin with?

What does each discrete sample represent?

What Does the Sample Represent?



What does a single sample represent and how can/should you assess the spatial variation of samples.

A small # of discrete samples encourages two key errors:

- 1) Underestimate the representative concentration in the an area, and
- 2) Underestimate the vertical and later extent of contamination.

Does 1g of soil represent your site?

That small volume of samples that is analyzed provides a results that represents the area we are assessing for risk type decisions.

Picture Reference: <http://www.swrcb.ca.gov/rwqcb2/brownfields.shtml>

What Do These Environmental Criteria Have In Common?



- ▶ Most risk-based environmental criteria based on estimate of mean
 - Soil screening levels
 - Regional screening levels
 - Site-specific cleanup levels
 - Exposure point concentrations

Most Risk-based environmental criteria are based on an estimate of the mean (.e.g., 95% UCL).

If you have a few discrete samples how do you estimate the mean?

Discrete samples give some sense of spatial variability.

The more discrete samples you have, the more you know about the population variability, and build certainty.

Uncertainty Sources



- ▶ Instrument analysis
- ▶ Sample preparation
- ▶ Laboratory sub-sampling
- ▶ Field sample collection



What are Sources of Uncertainty and where do we find them?

With proper maintenance and standard operating procedures instrument analysis is actually very consistent.

Uncertainty increases as you move down the list.

Uncertainty Sources



- ▶ Instrument analysis
- ▶ Sample preparation

- ▶ Laboratory sub-sampling
- ▶ Field sample collection



The largest amount of uncertainty lies in laboratory sub sampling and field sampling collection, including sample heterogeneity.

Is there a better way to control uncertainty and errors and get a better representative sample?

What is Incremental Sampling Methodology (ISM)?



ISM Objective: To obtain a single sample for analysis that has the mean analyte concentration representative of the decision unit

- ▶ Structured composite sampling and processing protocol
- ▶ Reduces data variability
- ▶ Provides a reasonably unbiased estimate of mean contaminant concentrations in a volume of soil targeted for sampling

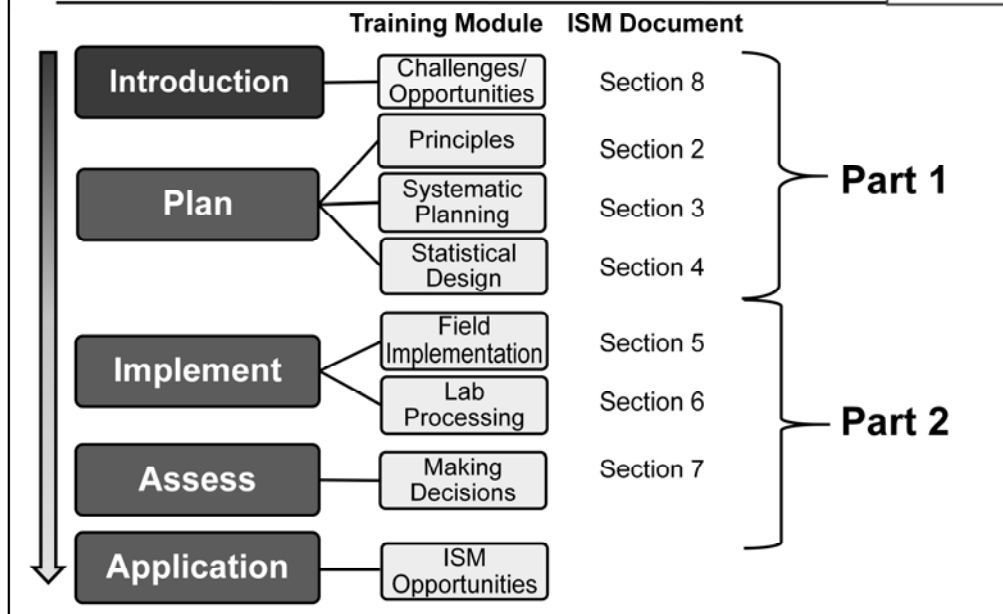
Decision Unit (DU): the smallest volume of soil (or other media) for which a decision will be made based upon ISM sampling

Incremental Sampling is a structured sampling and processing protocol that reduces data variability and increases sample representativeness.

ISM is an improved form of composite sampling, because the process involved in collection and analysis of an ISM sample greatly improves sample representativeness. The ISM sample goal is to have all the same constituents in the same proportion as the volume sampled.

ISM samples provide a result that better estimates the mean concentration of the sampled volume, than sparse discrete sampling.

ISM Document and Training Roadmap



The Internet-based training follows the ISM document section by section

Two-part ITRC Internet-based training:

Part 1 - Principles, Systematic Planning, and Statistical Design

Part 2 - Field Implementation, Lab processing, and Data Assessment

The planning, field implementation, and laboratory processing are critical to collecting a sample that yields Highly Reproducible Mean Concentrations

2009 ISM Survey: Areas of Question/Concern



263 responses (75% respondents state regulators and consultants)

- ▶ Can ISM find “hot spots”?
- ▶ Do regulators allow or accept ISM?
- ▶ Can you collect volatile organic compound (VOC) samples?
- ▶ Does ISM delineate the extent of contamination?
- ▶ What’s the right size for a Decision Units (DUs)?
- ▶ Can you obtain Upper Confidence Limits (UCLs)?
- ▶ Can ISM and discrete results be compared?
- ▶ Are there approved laboratory processes and certification?
- ▶ How much does ISM cost?

Reluctance to use ISM stems from a lack of experience

ITRC, ISM-1, Appendix B, August 2009 ISM Survey Results

Survey identified outstanding issues and used to aid development of the ISM technical regulatory document.

263 responses to our survey

½ of the responses from consultants

¼ from State regulators

¼ from federal agencies, regulators, laboratories, stakeholder groups.

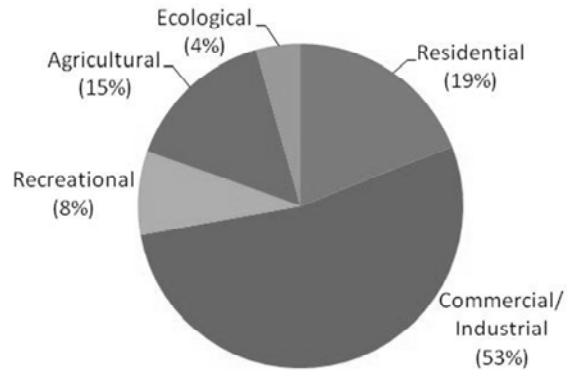
These summarize the key implementation issues identified by the survey.

Section 8 addressed these issues in the ISM document.

ISM is still being developed and some of these issues (e.g. cost) are still being worked out with more wide spread use.

2009 ISM Survey: ISM Sampling and Land Use

ISM primarily used at commercial/industrial sites
but applicable to all types of sites



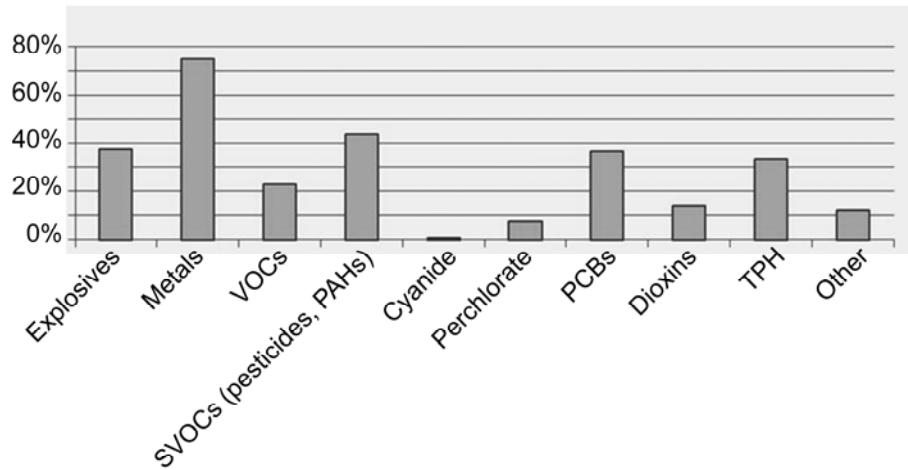
ITRC, ISM-1, Appendix B, August 2009 ISM Survey Results

The survey found that ISM has been applied on a variety of sites.

2009 ISM Survey: Chemicals of Interest for Incremental Sampling?



ISM can be used at sites with a broad range of contaminants



ITRC, ISM-1, Appendix B, August 2009 ISM Survey Results

ISM sampling can be used for a broad range of contaminants – data from our 2009 survey.

ISM – What is Your Perception? 2009 ISM Survey Revealed.....



- ▶ Common misperceptions: “It’s just composite sampling, misses hot spots, and costs more.”
- ▶ Few state regulators had heard of ISM and very few with ISM experience
- ▶ Hawaii, California, and Alaska made up over 40% of the reported ISM projects
- ▶ More than half of the state regulators responded that ISM was discouraged

We accepted the challenge to provide tools for state regulators, consultants, and others to learn the value of ISM and how to apply ISM

ITRC, ISM-1, Appendix B, August 2009 ISM Survey Results

No associated notes

ITRC ISM Team



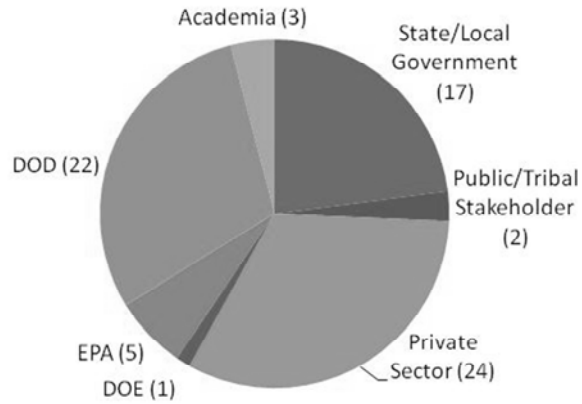
► Team history

- Formed 2009
- 74 members
- Regulators from 13 states

► Products

- Case study
- Web-based guidance
- Internet-based training

ITRC Team Member Composition



Multi-disciplinary team includes, scientists, geologist, toxicologists, engineers, chemists, statisticians, and community and tribal representatives.

19 Our ITRC Solution: ITRC ISM-1 Technical and Regulatory Guidance Document



Web-Based Document at:
<http://www.itrcweb.org/ISM-1/>



- ▶ Fundamental understanding of how and why ISM works
- ▶ Detailed instructions for design and implementation
- ▶ Addresses potential regulatory concerns
- ▶ Provides case studies and simulations



Training classes on Multi-Increment sampling and guidance available from Hawai'i, Alaska, and the USACOE

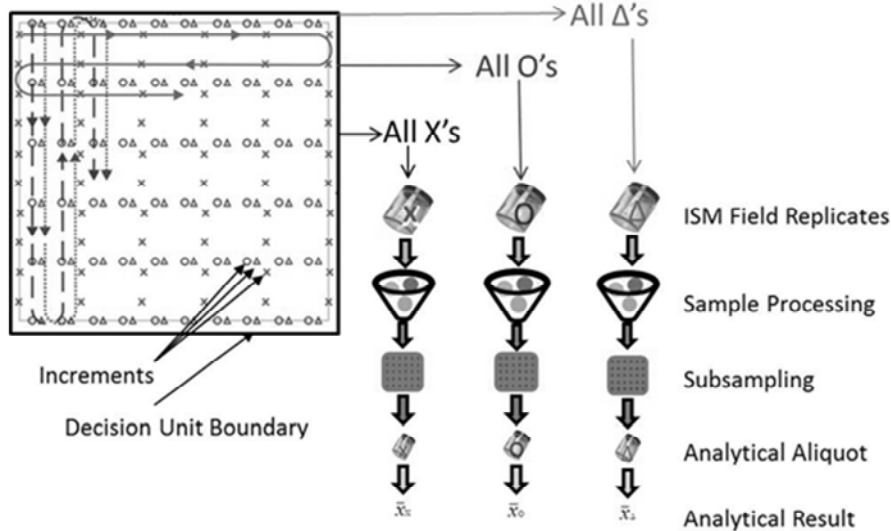
ISM Tech-Reg document deals with things in more detail and include case study and simulation info to support application of ISM.

In 2010, the ISM Team developed and implement an IS plan for a case study site in Florida. It is presented in our Tech-Reg document and will be presented during different Internet-based training modules.

As of March 2012, the Florida Dept. of Environmental Protection is in the process of developing ISM guidance by incorporating the ITRC ISM guidance document.

Web-based document has live links and you can print the entire document, a Section, or just a page.

What is Incremental Sampling Methodology (ISM)?



ISM involves planning, field implementation, and laboratory processing and subsampling to provide a representative analytical aliquot. The goal of ISM is to analyze an aliquot that represents the same proportion of constituents as the sampling area.

The box represents a Decision Unit or DU - the volume being sampled and the volume you want to make a decision on.

Within the DU, there are sample locations or "Increments".

The sampling grid and increment locations are established during the systematic planning as are the number of increments represented by the X's, O's and triangles.

The ISM field replicate for X is a composite of all 60 X increments, likewise for the O's and the triangles.

After collection in the field the samples are sent to a lab to be processed, and subsampled.

The lab subsampling approach is similar to that applied to collection of a single field increments. The goal of ISM is to have all the same constituents in the same proportion as the area sampled.

The ISM document recommends at least three replicate results (X, O, and triangle) for each DU.

Advantages and Limitations of ISM



Advantages of ISM	Effect
Improved spatial coverage (increments x replicates)	<ul style="list-style-type: none"> • Sample includes high and low concentrations in proper proportions
Higher Sample Mass	<ul style="list-style-type: none"> • Reduces errors associated with sample processing and analysis
Optimized processing	<ul style="list-style-type: none"> • Representative subsamples for analysis
Fewer non-detects	<ul style="list-style-type: none"> • Simplifies statistical analysis
More consistent data	<ul style="list-style-type: none"> • More confident decision

Limitations of ISM	Effect
Small number of replicates	<ul style="list-style-type: none"> • Limits Upper Confidence Limit calculation methods
No spatial resolution within Decision Unit	<ul style="list-style-type: none"> • Limits remediation options within Decision Unit • Limits multivariate comparisons
Assessing Acute Toxicity	<ul style="list-style-type: none"> • Decision Unit has to be very small

ISM has both advantages and disadvantages from a sampling design perspective. Can't directly compare discrete and ISM samples because each measure different properties of the population.

Under disadvantages, discrete sampling allows for calculations of ratios of two variables – allows for correlations among constituents, or estimates of bioaccumulation factors (update from abiotic media to organisms) that you cannot get from ISM.

When assessing acute toxicity issues, the decision unit would have to be very small for incremental sampling. ISM may not be practical.

ISM – What's In It For YOU?



- ▶ Fewer analyses but a more representative sample
- ▶ High quality data leads to a more confident decision
- ▶ Potential for cost savings

No associated notes

Are YOU the next ISM User?



This training will provide answers and show you how ISM fits with sampling and decision making.

How ? Why
? When ?
What Where

Training provides answers to make informed ISM decisions by answering:

Where can ISM be used?

When should ISM not be used?

What contaminants are most suitable for ISM?

What effect does sample processing have on contaminant concentration?

Does ISM mask area of high concentrations due to compositing and homogenization?

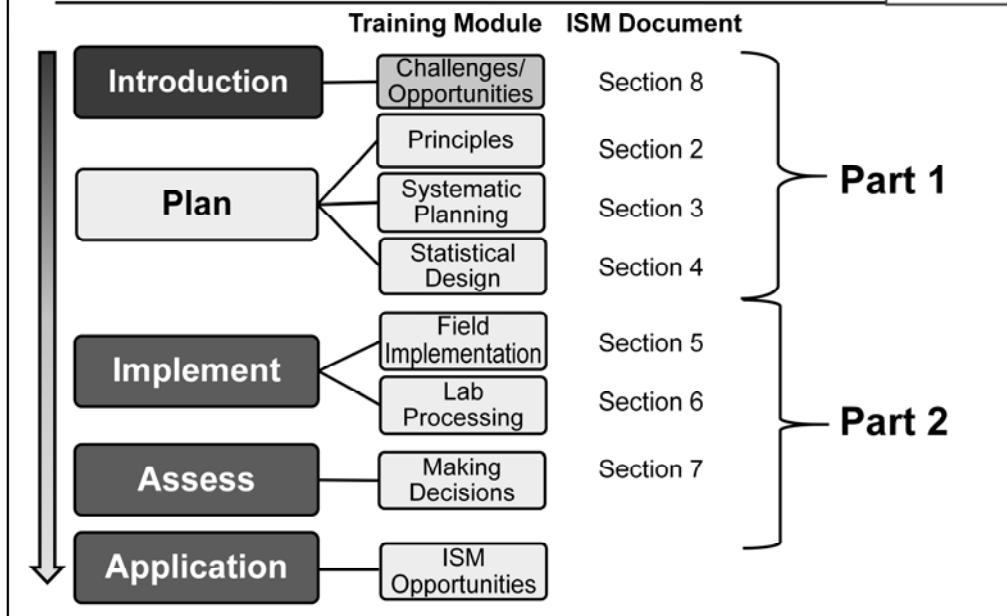
How does ISM differ from discrete sampling?

How many replicates should be collected?

How are data quality objectives (DQOs) addressed?

How do ISM results relate to action levels?

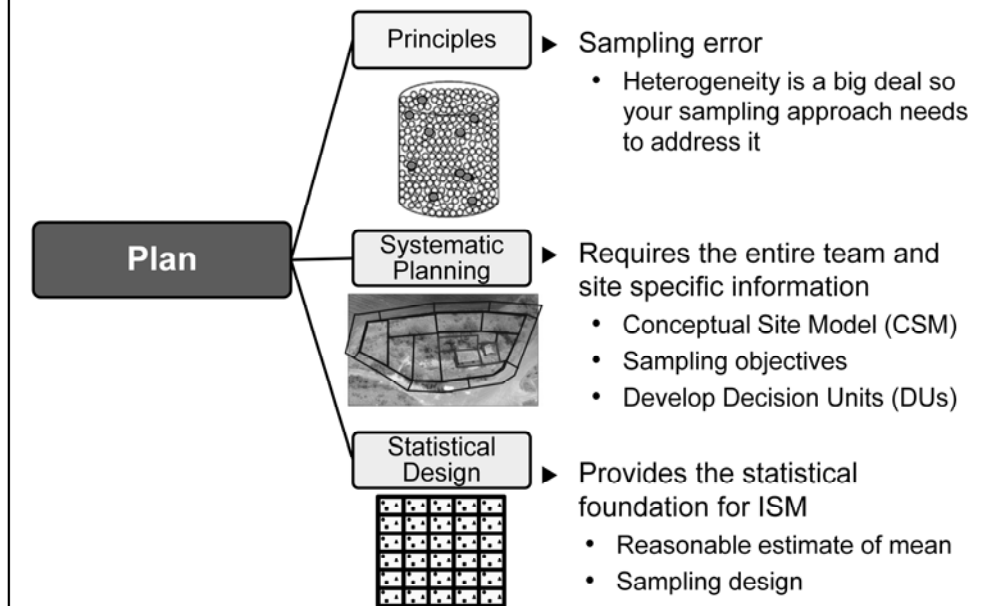
ISM Document and Training Roadmap



Part 1 includes Modules on:

Soil Principles, Systematic Planning, and Statistical Design

ISM Part 1 – Principles, Systematic Planning, and Statistical Design



Soil Principles – Be aware of issues related to heterogeneity and sampling errors

Systematic Planning - Involve the entire team, regulators, consultants, responsible parties in critical elements (e.g. conceptual site model, establish sampling objectives and decision units.) Sampling objectives should drive your sampling design, and the scale of decision making should align with sampling objectives.

Statistical Design - Provides the statistical foundation and describes why ISM provides a reasonable mean, describes a good ISM sampling design, and informs you how ISM provides 95% UCL.

Principles Learning Objectives



Learn how to use basic principles to improve planning, implementation and decision-making:

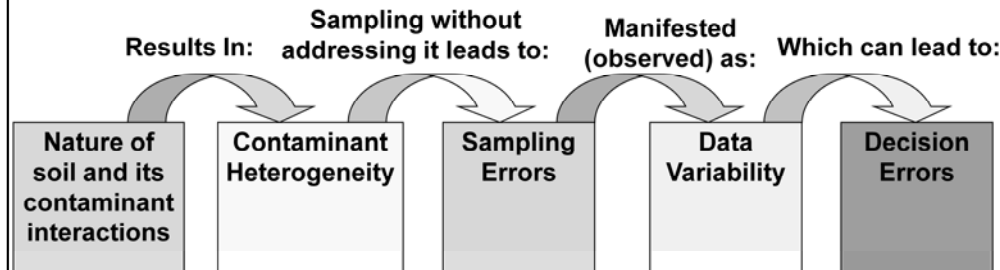
- ▶ Soil heterogeneity at 2 spatial scales makes it difficult to correctly interpret data results
 - Those spatial scales are micro-scale and short-scale
 - Heterogeneity at these scales can cause data variability → costly decision errors
- ▶ Micro-scale heterogeneity is managed by increasing sample mass and *improving lab sample processing (required by ISM)*
- ▶ Short-scale spatial heterogeneity is managed by the *field incremental sampling of ISM*

ITRC, ISM-1, Sections 2 and 5.3.1

Speaker Notes

- Micro-scale heterogeneity is also called compositional heterogeneity.

How Soil Heterogeneity Can Cause Decision Errors: Navigation Pane



- ▶ Heterogeneity: the condition of being non-uniform
- ▶ The heterogeneous nature of contaminants in soils increases the chances of decision error

ITRC, ISM-1, Section 2.1

Speaker Notes

Previous section emphasized 1) importance of up-front planning BEFORE going to the field, and 2) sample processing is critical part of incremental sampling methodology. Planning is necessary to minimize errors throughout data generation. In order to plan effectively, need to understand the causes and effects of sampling error, and be able to detect when sampling error may have degraded data quality.

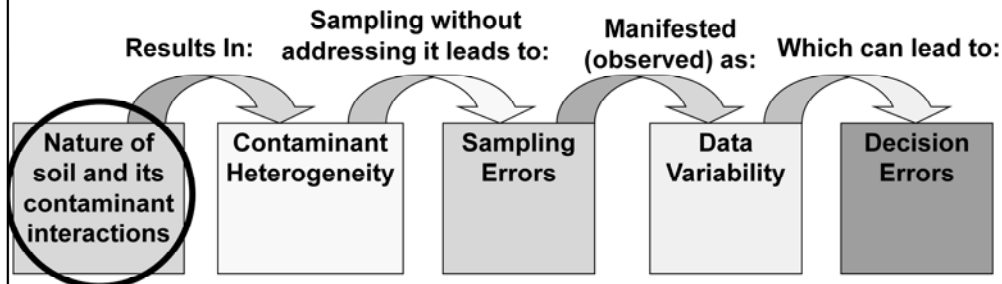
A key goal of ISM is to reduce decision errors when dealing with soils. Decision error refers to decisions that would be made differently IF the true nature of the contamination were known. An example of a decision error is deciding that contamination is not present above a certain level, when it actually is.

To reduce decision errors, need to examine their root causes. This slide shows how the heterogeneous nature of soil ends up causing decision errors. Starting with the first box on the left (the blue box), we'll look at how contaminants interact with soil. That interaction leads to contaminant heterogeneity, which is a primary cause of sampling errors. Sampling errors, in turn, lead to data variability, and data variability can mislead decision makers. In this presentation I'm going to explain how this cascade occurs. We'll use this figure to serve as a navigation aid.

Supplemental Information

See ISM-1 Section 2.1

Soil is a Complex Particulate Material



- ▶ All soil is heterogeneous in composition
- ▶ Typical mixing/stirring cannot make soil uniform

ITRC, ISM-1, Section 2.2

Speaker Notes

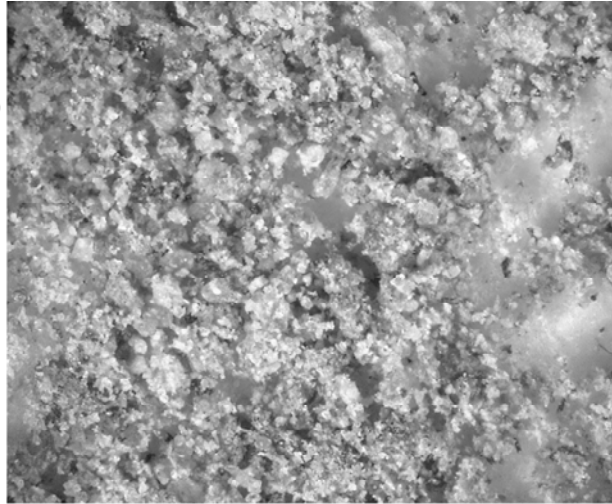
- Heterogeneity in composition is referred to as “compositional” or “constitutional” heterogeneity
- Soil samples show compositional heterogeneity in 3 primary ways:
 - made up of particle sizes that vary over several orders of magnitude and so differ in surface area, and
 - composed of different mineral grains which vary in their “stickiness” for contaminant molecules,
 - contain various amounts and types of organic matter.

Supplemental Information

See ISM-1 Section 2.2

Micro-Scale Variation in a Homogeneous-Looking Soil

Photo credit: Deana Crumbling



A sandy soil, showing variation in particulate size and mineral content (10X magnification)

Speaker Notes

- This figure is a photo of magnified sandy soil containing very little organic matter.
- At the macro level (i.e., viewed without magnification) this soil looks homogeneous.
- But under magnification, can see it is composed of mineral grains that vary from relatively large to barely visible at this magnification.
- Colors of individual grains vary from white to pink to greenish, reflecting the different minerals present in the grains.
- This is an example of soil heterogeneity at the micro-scale.

Soil Particle Composition



Individual soil particles are inorganic mineral or some form of organic carbon.

- ▶ Many contaminants adhere to the surfaces of certain minerals
- ▶ Organic carbon is composed of complex molecules that act as molecular sponges

ITRC, ISM-1, Section 2.2

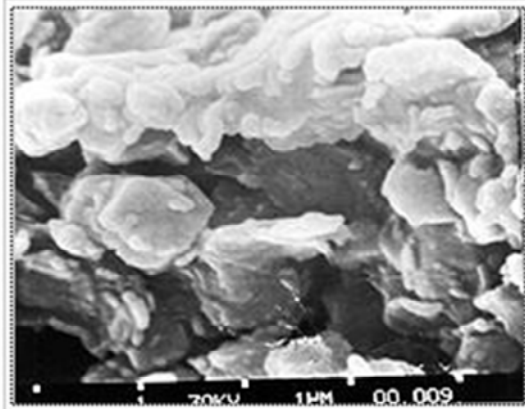
Speaker Notes

- Soil is composed of an infinite number of particles, some made of inorganic minerals and some of organic matter.
- Organic matter comes from living organisms. It could be grass, leaves, sticks, insects, microbes, etc.
- Organic matter is of particular importance to contaminants because it decays into complex molecules that act as molecular sponges.
- Many contaminants, although not all, adhere well to certain soil minerals.
- Organic and inorganic contaminants can be absorbed into organic carbon complexes.

Supplemental Information

See ISM-1 Section 2.2

“Sticky” Minerals



Electron microscope photograph of smectite clay – magnification 23,500

- ▶ Contaminant molecules/atoms “stick” well to certain particles
- ▶ Smallest particles usually the stickiest
 - Clays (see photo)
 - Iron (hydr)oxides
- ▶ Stickiness mechanisms
 - (-) and (+) charges
 - Surface area

ITRC, ISM-1, Section 2.2.1.1

Photo credit: USGS, 2006

Speaker Notes

- Different particles bind contaminants to varying degrees.
- Clay minerals strongly bind many contaminants.
- This is a photograph of clay particles under high magnification. If you look closely, you’ll see that clay particles take the form of stacked plates.
- The plates have molecule-sized spaces between them. The very small size of clay particle gives them a large surface area on the outside, then the plate structure provides even more. More surface area provides more sites where contaminants can bind.
- Some of the “stickiness” of clays is due having many negative charges lining the inside of the plates. Clay particles have some positive charges along the edges of the plates.
- Negative charges attract metal contaminants that are positively charged. An example is when Pb in bullets corrodes into Pb minerals that then dissolve and release positively charged Pb ions into the soil. Lead ions are attracted to the clay’s negative charges and can become trapped between the clay plates.
- Another type of soil particle that is sticky from a contaminant’s point of view are oxide minerals, such as iron oxides and aluminum oxides. Geochemical oxides are small particles with a large surface area. They can carry either a positive or negative charge depending on the pH.
- Iron oxide is interchangeable with iron hydroxide, depending on the pH.

Supplemental Information

See ISM-1 Section 2.2.1.1

Photo credit: USGS Photo Library, 2006, USGS, URL = http://libraryphoto.cr.usgs.gov/cgi-bin/show_picture.cgi?ID=ID.%20McKee,%20E.D.%20%20316

Particles with High Loadings are Called “Nuggets”

- ▶ Contaminants adsorbed to distinct particles form “nuggets” of high concentration

“the iron in a cubic yard of soil [1-1.5 tons] is capable of adsorbing 0.5 to 5 lbs of soluble metals ...or organics” (Vance 1994).

Arsenic (whitish color) sorbed to iron hydroxide particles

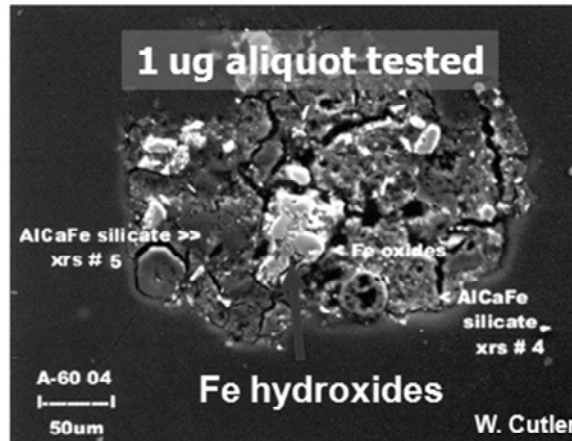


Photo courtesy of Roger Brewer, HDOH

ITRC, ISM-1, Section 2.2 hyperlinks

Speaker Notes

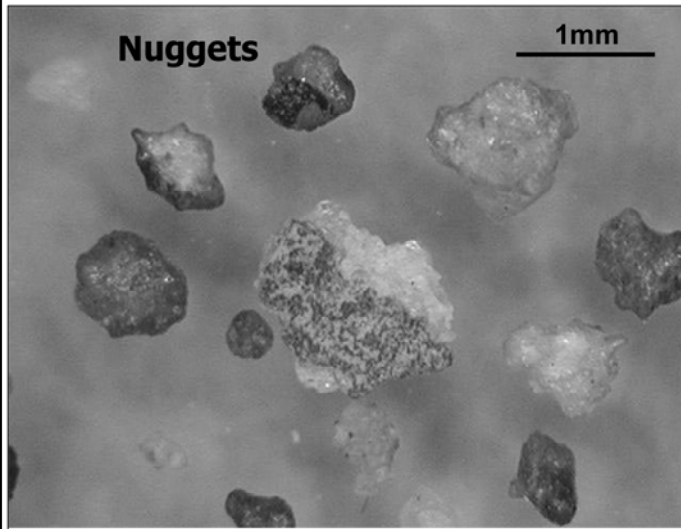
- Particulate iron minerals, such as oxides, are very good at binding contaminants.
- One researcher stated that the Fe in a cubic yard of soil can adsorb ½ to 5 lbs of soluble metals or organics.
- The photomicrograph shows microscopic iron hydroxide grains coated with arsenic. The arsenic appears as a light-colored deposit covering Fe-OH grains (see red arrow).
- Silicate minerals make up most of the soil mass in the photo. Arsenic does not adsorb to those minerals, so they are dark gray.
- Photo provided by Roger Brewer with the Hawaii Dept. of Health

Supplemental Information

See ISM-1 Section 2.2 hyperlinks

Quote from the journal article: “Given the average concentration in soil, the iron in a cubic yard of soil is capable of adsorbing from 0.5 to 5 pounds of soluble metals as cations, anionic complexes, or a similar amount of organic[s].” (Vance, 1994). [Reference = David B. Vance. “Iron – The Environmental Impact of a Universal Element,” *National Environmental Journal*, May/June. 1994 Vol.4 No. 3 page 24-25. see also URL = <http://2the4.net/iron.htm>]

Key Point: Contaminants Often Exist or Behave as Particles



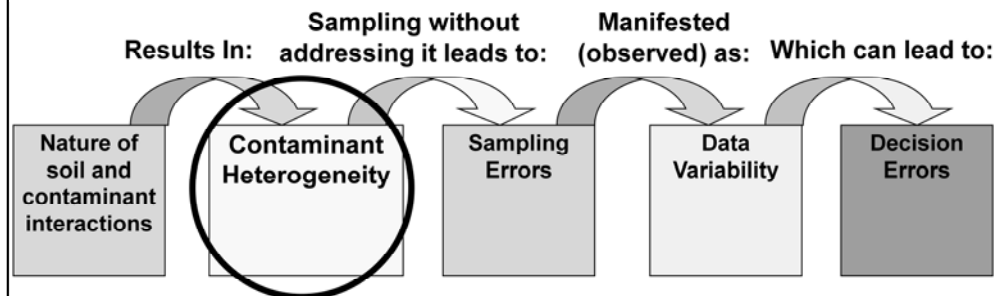
Tiny chunks of pure TNT-based explosive compound isolated from a soil sample

Photo courtesy of Alan Hewitt (USACE)

Speaker Notes

- The photo is of Composition B particles from low-order 81-mm mortar. 60% Military grade RDX (Contains about 10% HMX) 39% Military grade TNT.
- Photo provided by Alan Hewitt (US Army Corp of Engineers Cold Regions Research and Engineering Laboratory).
- Sometimes contaminants are released directly in particulate form. Examples are explosives residues, organic or metal-based pesticides applied as a dust; airborne smelter residues depositing as dust; and lead and other metals dust and fragments created by firing guns at firing ranges.
- But even if they were not originally released in particulate form, contaminants behave as if they were particles when they bind to soil particles by the mechanisms just described.
- As a consequence contaminants are heterogeneous in their spatial distribution throughout even small soil samples.

Particulates in Solid Matrices Create “Micro-Heterogeneity”



- ▶ “Micro-heterogeneity” is non-uniformity within the sample jar
- ▶ Important because contamination is heterogeneous **at the same spatial scale as sample analysis**

ITRC, ISM-1, Section 2.5.2

Speaker Notes

- This is important because particulate contaminants integrate with soil particles in a non-uniform manner, which creates contaminant heterogeneity at a micro-scale. In other words, contaminants are not uniformly spread out evenly throughout the soil in a jar.
- This is called “distributional heterogeneity” at a micro-scale.
- This matters because the mechanics of sample analysis take place at this micro spatial scale. When the lab scoops out subsamples from a jar for analysis, different scoops of soil may have different numbers of contaminant particles.

Supplemental Information

See ISM-1 Section 2.5.2

Micro-Heterogeneity Makes Contamination Hard to “Read”



- ▶ Micro-heterogeneity interferes with interpreting analytical results
- ▶ If contaminant distribution is not uniform in the sample jar, how sure that analytical data represent the contents of the jar, much less the field?
 - Huge mismatch between scale of decision-making and scale of sample analysis

ITRC, ISM-1, Section 2.4

Speaker Notes

- Like in this graphic, heterogeneity makes soil contamination hard to read.
- Soil may appear to be homogeneous when viewed from the spatial scale of decision-making, but it is NOT homogeneous at the scale at which chemical data are generated.
- Yet we expect that analyzing tiny samples will tell us the true concentration of tons of soil.

Supplemental Information

See ISM-1 Section 2.4

Metals Analysis on 1 Gram of Soil Guides Decisions on Tons



VS.

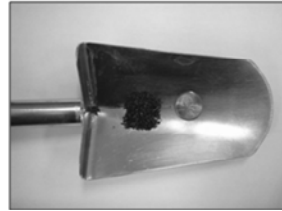


Photo credits:
Roger Brewer, HDOH

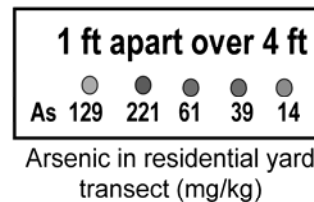
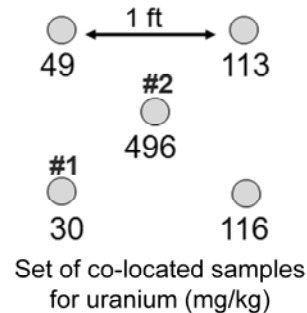
Speaker Notes

- Based on the results of analyses performed on a few grams of soil, decisions are made about whether contamination is present (and at what level) in tens to hundreds to thousands of tons of soil.
- Although a jar of soil containing 100 or more grams of soil is submitted to the lab, routine metals analysis actually analyzes only 0.5, 1 or sometimes 2 gram of soil (depending on the lab) from that jar.
- Organics analysis typically will analyze from 5 to 30 grams (depending on the lab and the analyte).

Short-Scale Field Heterogeneity: Co-located Samples



- ▶ Shortest spatial scale in the field measured by “co-located samples” (inches to a few feet apart)
- ▶ Samples anticipated to be “equivalent,” but often give very different results
- ▶ Chance governs exact location where soil is scooped
 - Therefore, **chance** can determine decision outcome!
- ▶ *ISM addresses the problems of both micro- and short-scale heterogeneity*



ITRC, ISM-1, Section 2.2.2

Speaker Notes

Previous discussion focused on heterogeneity WITHIN a sample (i.e., within a single jar).

Now will focus on heterogeneity BETWEEN samples (i.e., from one sample to another in the field).

“Co-located samples” are a QC check designed to assess short-scale heterogeneity.

“Co-located samples” are expected to be equivalent in that they are expected to have pretty much the same concentration.

But often co-located samples have very different results even though they may be only inches apart.

Within the confines of a small area, chance governs which spoonful of core of soil is picked to be put in the jar.

Since the concentration of each spoonful might be very different, chance can determine which decision gets made on that area.

This is one of the dangers of making decisions based on single grab sample results.

Uranium data set: although only about 8 inches apart, sample #1 has a concentration of 30, while sample #2 is nearly 500. If discrete samples are used to make decisions, the decisions in this area would be determined by chance, because the result depends on where the sampler happens to kneel down and dig.

Co-located sample results are affected by both short-scale heterogeneity AND within-sample heterogeneity. So unless you have controlled for within-sample heterogeneity, you won't be able to measure between-sample heterogeneity.

Uranium data source: Robert Johnson (US Dept of Energy)

Arsenic data source: Deana Crumbling (USEPA). Data from center of a residential yard.

Supplemental Information

See ISM-1 Section 2.2.2

Long-Scale Heterogeneity is Generally at the Scale of Decision-Making



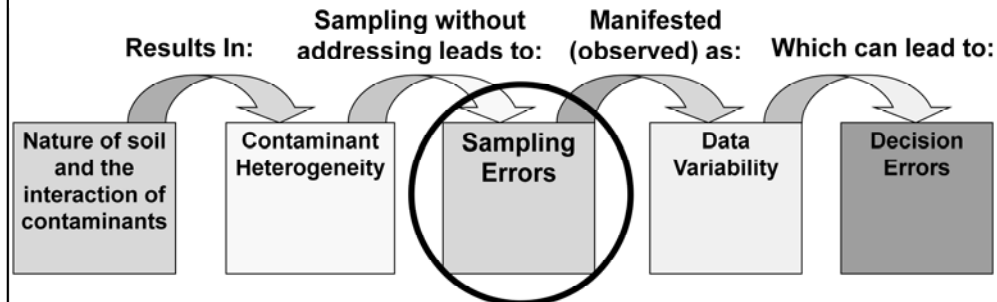
Figure credit: Roger Brewer, HDOH

Results for an actual sampled property. Green circles denote concentrations below the action level; red circles are above the action level.

Speaker Notes

- Long-scale heterogeneity is the spatial scale at which differences in concentration are expected.
- Sampling programs are generally designed to search for variation in concentrations at this scale.

Heterogeneity Causes Sampling Errors



- ▶ Sampling error occurs when samples fail to mirror (represent) the original targeted population
- ▶ Need the concept of “sample support” (the physical dimensions and mass of the sample)

ITRC, ISM-1, Section 2.3.2, 2.4.1.1 and 2.2 hyperlinks

Speaker Notes

- The next concept is that insufficient control over heterogeneity’s effects can lead to sampling errors. A sampling error is said to occur when the sampling process produces a sample that does not represent the intended population.
- To discuss sampling errors, need the term, “sample support.”
- Sample Support: “the size, shape, and orientation of sampling volume (i.e., “support”) for heterogeneous media have a significant effect on reported measurement values.” EPA Soil Screening Guidance: Technical Background Document, EPA/540/R95/128, May 1996
- Sample support refers to the physical dimensions of a sample as it is collected from the parent material.
- For example, a 6-inch deep core is a different sample support than scraping up surface soil to a 2-inch depth. A core with a 2-inch diameter is a different sample support from a core with a 5-inch diameter. A 100-gram sample is a different sample support from a 300-gram sample.
- Sample support is a critical factor governing the results of soil measurements.
- The following slides explain how this works.

Supplemental Information

See ISM-1 Sections 2.3.2, 2.4.1.1 and 2.2 hyperlinks

Concentration is a Function of Sample Support and Nugget Mass

Common assumption

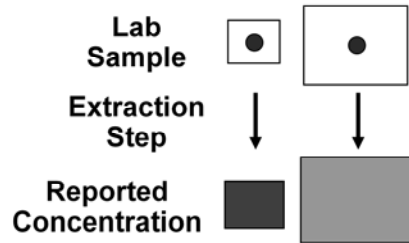
The amount of soil analyzed makes no difference to what results are obtained.



Concentration (mg/kg) =
contaminant mass (mg) /
the soil mass (kg)

Assumption wrong for solids

Can have the **same** contaminant nugget mass (blue), BUT in different **sample** masses (white)...

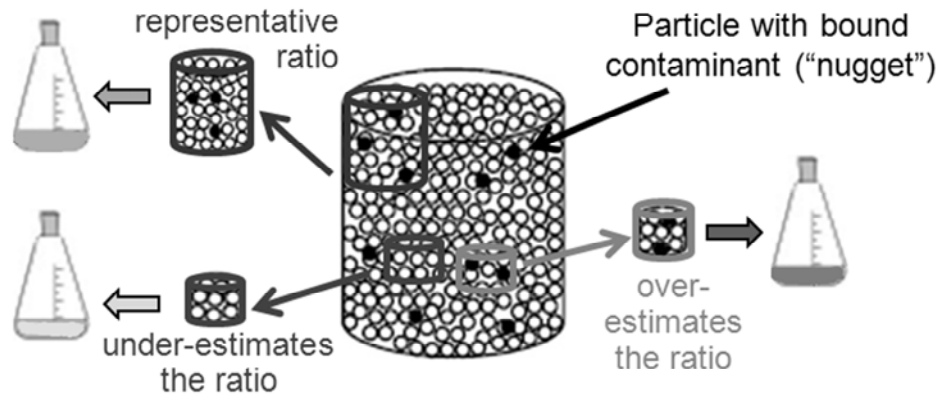


...get **different** concentration results

Speaker Notes

- The issue of “sample support” for heterogeneous environmental and waste matrices invalidates the common assumption that the reported concentration of an environmental sample should be the same no matter what mass/volume of sample is collected and analyzed.
- The mass/volume of the sample greatly influences the reported concentration for a sample, especially when contaminants are heterogeneously distributed throughout the parent matrix.
- It is like putting a drop of dye in water in a water glass vs. in water in gallon jug. The water in the glass will have a more intense color than the water in the jug.
- For heterogeneous samples (which are affected by the nugget effect to a greater or lesser degree), the analytical result for a sample is determined by how much contaminant (in the form of concentrated nuggets) is captured in that sample amidst a volume of cleaner matrix.
- The cleaner matrix serves to “dilute” the concentrated particles during sample extraction (for organics) or digestion (for metals).
- The issue of sample support is becoming an increasingly important determinant of analytical results as more sophisticated analytical technologies and efforts to reduce generation of lab waste drives a trend toward smaller and smaller masses of sample.

Smaller Sample Supports More Prone to Sampling Error than Larger Ones



- ▶ Illustration of sampling error: For the blue and green samples, the proportion of nuggets in the samples do not represent the nugget proportion of the population (the large container)

Speaker Notes

- Another way sample support influences concentration results is whether the sample support is large enough to accurately capture the particle ratio of the population.
- In this cartoon, the large container represents a field sample in a jar.
- The cartoon illustrates how subsample support affects how well a lab subsample represents the field sample.
- The lab subsample represents the field sample if the ratio of contaminant-laden particles to "cleaner" particles in the subsample mirrors the ratio in the field sample.
- A sampling error occurs when a subsample does not have same ratio as the field sample.
- Sampling error is more likely for smaller subsample supports, since they are more likely to under- or over-estimate the proportion of "hot nuggets" to less contaminated "cool" particles.
- Larger supports are more likely to represent the actual ratio and give a concentration result that is representative of the mean of the jarred field sample.
- Figure adapted from EPA 530-D-02-002, RCRA Waste Sampling Draft Technical Guidance, August 2002, page 92.
- ISM addresses this problem by collecting many increments which results in a large mass representing the whole volume of the material being investigated.

Change the Sample Support and Change the Concentration

Concentration (mg/kg) =
contaminant mass (mg) /
the soil mass (kg)

Arsenic mass of 5 ng in a
sample support of 1 μ g of
other soil minerals: arsenic
conc = 5000 mg/kg

Analyze an As-Fe-OH grain
by itself and arsenic conc
might be 100,000 mg/kg
(10%) or more.

Arsenic (As) sorbed to iron hydroxide
(Fe-OH) mineral grains

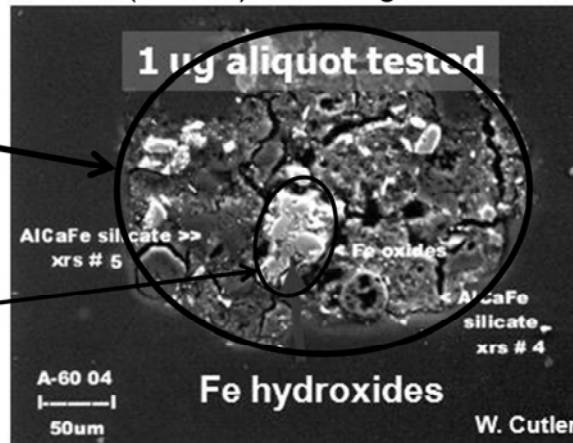


Figure courtesy of Roger Brewer

Speaker Notes

- Nuggets carrying high contaminant loading have a huge effect on what the concentration is reported to be. Concentration is determined by 2 things: the mass of the contaminant and the mass of the material that contains the contaminant.
- A smaller mass of soil that contains some contaminant-laden nuggets will have a higher concentration than if the same nuggets are present in a larger mass of soil.
- In the picture, the mass of arsenic on the Fe-OH mineral grains was measured to be 5 nanograms. The mass of the soil minerals containing the arsenic is 1 microgram (blue circle). Expressed in common concentration units, 5 ng arsenic in 1 microgram of soil material is 5000 mg/kg.
- On the other hand, consider if only the arsenic-coated iron hydroxide particle itself were analyzed (red oval). The arsenic might make up 10% of the mass, while iron hydroxide makes up the rest (90%). Then the arsenic concentration would be 100,000 mg arsenic/kg of soil material.
- When the number of contaminated particles (i.e., the mass of contaminant) stays the same, the concentration will be different depending on how much soil material is digested and analyzed.
- *The notion of “maximum concentration” is meaningless unless a sample support is specified. This applies in the field as well as in the laboratory.*

ISM Addresses Sample Support

Same As-Fe-OH grains
in 1 gram of other
minerals: arsenic
conc = 0.005 mg/kg



Photo credit: Deana Crumbling

A lack of control over sample support during lab subsampling and in the field is a primary cause of sampling error and data variability.

ISM explicitly manages sample support!

ITRC, ISM-1, Sections 5 and 6

Speaker Notes

- But what if those same arsenic-bearing grains were present in 1 gram of “cleaner” soil particles (particles that are not laden with arsenic)?
- Then the arsenic concentration would be 5 ng arsenic in 1 gram of soil or 0.005 mg/kg.
- A key concern of ISM is controlling sample support!

Supplemental Information

See ISM-1 Chapters 5 and 6

Ways to Reduce Sampling Error When Sampling a Jar

- ▶ ISM stresses the importance of sample support and techniques to reduce sampling error
 - Reduce particle size (grinding)
 - Increase sample support (i.e., extract a larger analytical sample mass)
 - Take many increments to make up the analytical subsample (“incremental subsampling”)
 - Use equipment like rotary splitters →



ITRC, ISM-1, Table 3-1 and 6.2.2.5 to 6.2.2.7

Speaker Notes

- Unlike routine discrete sampling programs, ISM specifically addresses sample support issues. A project team using ISM must consider the likelihood of nuggets, the analytical subsample’s volume and particle size.
- Reducing the overall particle size by grinding prior to subsampling may sometimes be required.
- Increasing the mass of the subsample and incremental subsampling are common ways to reduce subsampling error.
- If a field sample needs to be split, there are specialized equipment and techniques, such as rotary splitters. Choice of technique is heavily dependent on soil properties.

Supplemental Information

See ISM-1 Chapter 6

See also EPA guidance documents:

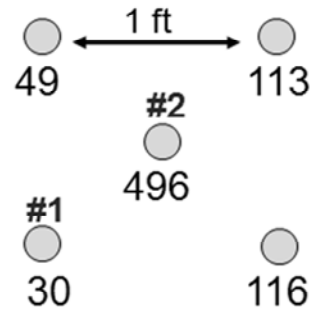
- “Guidance for Obtaining Representative Laboratory Analytical Subsamples from Particulate Laboratory Samples”, EPA/600/R-03/027 (Nov 2003); and
- “RCRA Waste Sampling Draft Technical Guidance”, EPA 530-D-02-002 (August 2002), Chapter 6

Reducing Short-scale Sampling Error



► Goal is to get THE concentration for a target soil volume, so...

- IDEAL: analyze whole volume as a single sample
- PRACTICAL: Increase sample support and sampling coverage by taking many small increments across the area and pooling them



► *This is what ISM does*

Set of co-located samples for uranium

ITRC, ISM-1, Section 2.6.2.1

Speaker Notes

The same principles apply to short-scale sampling error. Recall that this refers to extrapolating a single data point to a large field area without taking spatial heterogeneity into account.

Taking the whole targeted soil volume as a single sample for analysis would provide THE concentration for that volume without any sampling error. But, of course, that is not possible, so we take samples.

Need to have enough samples to include fluctuations in concentration in the result for the soil volume, but without exorbitant cost.

This goal can be accomplished by taking increments from many locations and pooling them together for a single analysis.

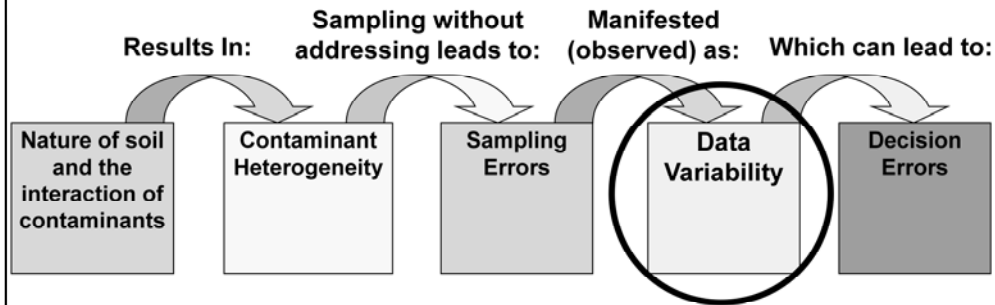
Incremental field sampling increases the sampling density (the number of samples per unit area) AND it increases the sample support of the field sample—both of which help control sampling error.

This is what ISM does in its planning stage and field implementation stage.

Supplemental Information

See ISM-1 Section 2.6.2.1

Sampling Error Causes Data Variability



► Sampling errors contribute to data variability

ITRC, ISM-1, Sections 2.4.1.3

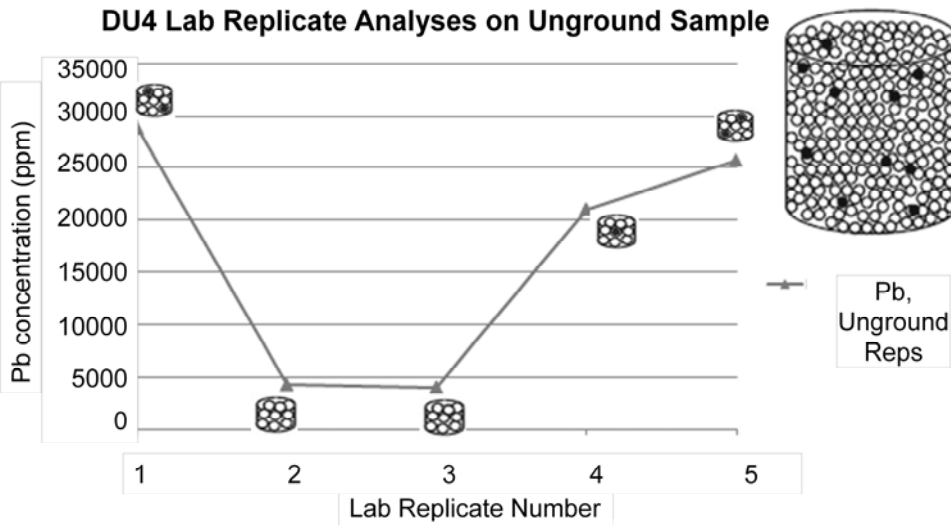
Speaker Notes

- Using our navigation sequence, we see that sampling errors are commonly observed as data variability.

Supplemental Information

See ISM-1 Sections 2.4.1.3

Study Data for Pb: 5 Laboratory Replicate Subsamples from Same Jar



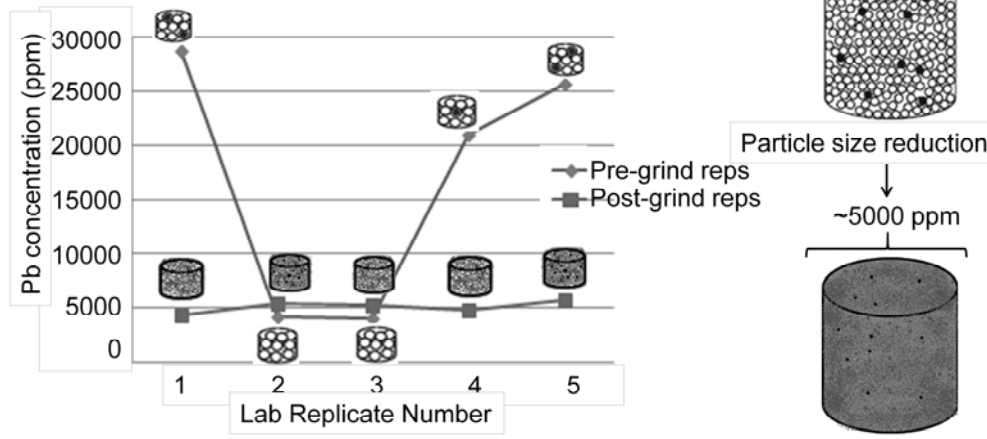
Speaker Notes

- Don't dismiss this data just because all the concentrations are high. Just because these Pb concentrations are much greater than the common risk-related threshold of 400 ppm does not mean that variability at these high concentrations are not important. Decisions about remedy selection and design or soil treatment and disposal may still hinge on differences at these high concentrations.
- The prime purpose of this graph is to illustrate the extreme variability that soil contamination can display.
- Soils that are contaminated are more likely to display a nugget effect which manifests as high variability.
- Soils that are not contaminated (or very lightly contaminated) are less likely to have particles with high contaminant loading, and so typically show less variability.
- Data variability is striking in this experiment where 5 replicate subsamples were taken from a single unground sample. Each of the 5 subsamples were analyzed for metals. The mass of the subsamples was 2.5 grams.
- The Pb results varied between 4000 and 28,000. Remember! These are not different field samples...they are 5 different subsamples from the same jar of soil.
- Fortunately, routine lab quality control checks provide measures of variability. QC includes co-located samples, field splits, lab duplicates, and matrix spike/matrix spike duplicates.
- Unfortunately, the information provided by these QC results is greatly under-appreciated and often ignored.
- A small sample mass composed of large particles frequently does not preserve the proportion of constituents as is present in the original population.

Same Soil Sample After Grinding

Pre-grind range: Pb 4000-29000 Post-grind range: Pb 4360-5660

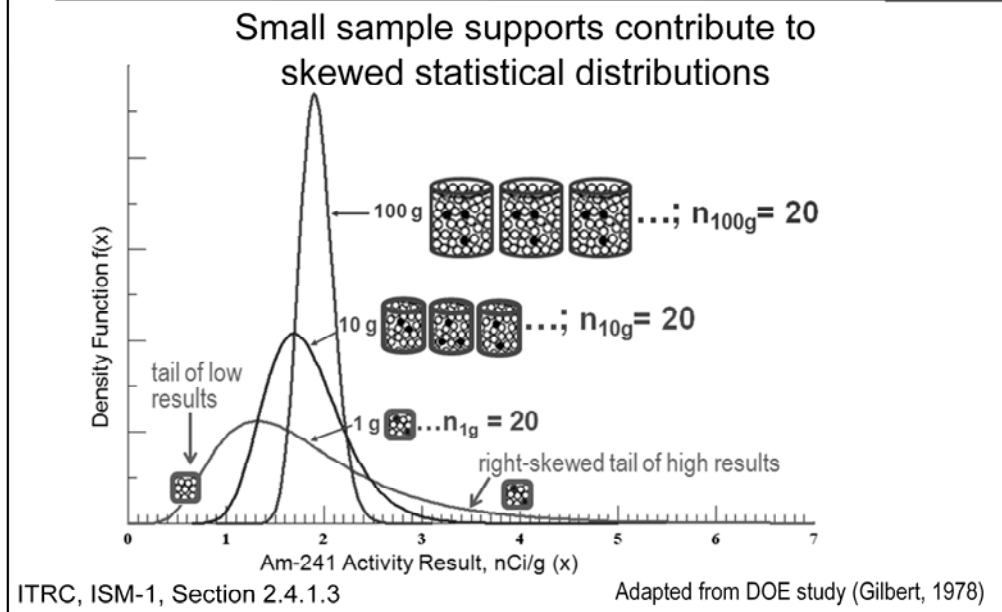
DU4 Pb Unground vs. Ground Subsample Replicate



Speaker Notes

- This is a continuation of the previous slide, now with ground results for the same sample ("post-grind", in pink) also included in the graph.
- 5 replicate subsamples were taken for analysis after the sample had been ground. The mass of the subsamples was again 2.5 grams.
- Variability was markedly reduced, which is the same as saying precision was markedly increased.
- The dramatic influence of nugget effects in the unground sample is evident by comparing the 2 sets of replicates.
- This data illustrates how grinding provides the smaller particles and mixing needed to better preserve the sample's constituent proportions even when small subsamples are used.
- The larger the particle size in the sample, the more subsample mass is needed to produce a representative subsample.

Sample Support Influences Statistical Distributions



Speaker Notes

- This graph plots the data from a study done in the 1970s. It directly measured how different masses of analytical samples (i.e., the sample support) influenced the statistical distribution of the data.
- Measurement units are in nCi/g, a measure of radioactive activity, which is related to the concentration (in ppm) of the radioisotope, in this case americium-241.
- The experiment involved first preparing a large soil sample (with mass of several kilograms—not shown on the slide—which will be called a batch) from which subsamples of various sizes could be taken (as shown on the slide). Preparing the large batch involved moderate homogenization efforts involving mild grinding and then sieving to less than 10-mesh.
- A series of 20 subsamples each of different supports were taken from the large prepared batch.
- The subsample supports that were tested included 1-gram, 10-gram, and 100-g ram.
- The wider the peak shape, the more variability present in the data set.
- The data set from the 1-g subsamples plots as a statistical distribution that is unsymmetrical and skewed in that the right-hand tail is pulled out.
 - The 1-g tail does not reach the x-axis until about 5 (note the green subsample on the right with a higher nugget:matrix ratio than the ratio in the 100-g samples).
 - Many samples have low concentrations, reaching down to about 0.25 (green subsample on the left without any high-load nuggets)
 - The width and shape (a low hump) of the curve mean that repeated 1-gram subsamplings of the large batch will produce data results that have a wide spread in values. Frequently there are low results, but sometimes there will be very high results. This variability is also called imprecision. No single result can be trusted to be close to the true mean of the batch.
- In contrast to the 1-g subsamples, the 20 10-g subsamples (purple) showed much less skewing of the right tail.
 - The right-hand tail reaches the x-axis just past 3.
 - The left-hand tail shows fewer samples (than the 1-g data set) with very low results, with the lower range of the distribution ending at about 0.8
 - The width of the 10-g peak is narrower, reflecting less variability (more precision) in the 10-g data set
- For the 100-g subsamples (red), the statistical distribution is almost symmetrical, with a high tight peak (high precision) and the right skewing nearly gone.
 - The 100-g curve reaches the x-axis on the right at about 2.5
 - On the left, the 100-g curve runs only down to about 1.4
 - The height and narrowness of the 100-g peak indicates that replicate subsamplings of the batch produce values that are close to each other (precise), and most likely close to the true mean for the large batch.
- Not only do small sample supports increase variability, they also contribute to data taking a lognormal or gamma (or other skewed) statistical distribution.
- So what does this have to do with decision errors?

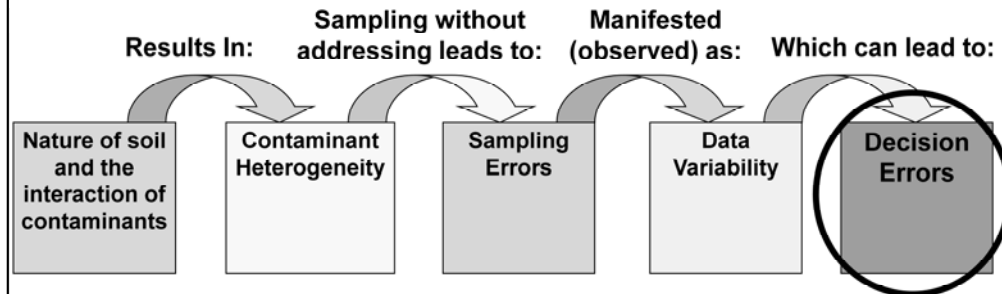
P.G. Doctor and R.O. Gilbert. 1978. DOE NAEG Report. *Two Studies in Variability for Soil Concentrations: with Aliquot Size and with Distance* [provided in webinar References]

See also Gilbert, Richard O. and Pamela G. Doctor. 1985. *Determining the Number and Size of Soil Aliquots for Assessing Particulate Contaminant Concentrations*. Journal of Environmental Quality Vol 14, No 2, pp. 286-292.

Supplemental Information

See ISM-1 Section 2.4.1.3

Concepts Underlying ISM: Avoiding Decision Error



- ▶ Decision Error: a decision that would have been made differently if the true condition were known
- ▶ Can occur when conclusions are based on data that were significantly influenced by heterogeneity

ITRC, ISM-1, Section 2.4.1.3 and 2.4.2

Speaker Notes

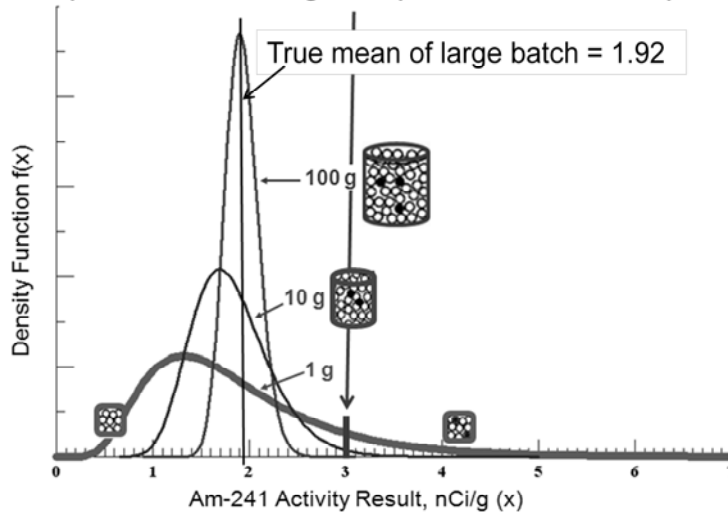
- A decision error is a conclusion that is different from the conclusion the data user **WOULD** have made if the true condition were known.
- Decision error in this context refers to using data results to draw a conclusion without taking data variability and other sources of sampling and analytical error into account.

Supplemental Information

See ISM-1 Sections 2.4.1.3 and 2.4.2

Skewed Data Distributions Promote Decision Errors

Suppose 3 is an action level. The likelihood of single data points exceeding 3 depends on the sample support.



Speaker Notes

- It is known that the true “concentration” of the large multi-kg batch is 1.92 (refer back to a previous slide)
- Measurement units are in nCi/g, a measure of radioactive activity, which is related to the concentration (in ppm) of the radioisotope, in this case americium-241.
- For the sake of discussion, suppose 3 is an action level, which is here shown as the small vertical blue line on the x-axis.
- Therefore, the true “concentration” of the large batch (1.92) is below the action level of 3
- The question for a data user is: Will the subsample that is analyzed lead to the correct conclusion about whether the “concentration” of the batch is higher or lower than 3, or could the data lead the user astray?
- Look again at the curve representing the 1-g subsamples (the heavy-lined curve): Even though the true mean is well below 3, the skewed nature of the data means that sometimes (around 11% of the time) data results are going to be higher than 3, as exemplified by the green subsample on the right. This would contribute to a decision error.
- Note that there is a 12% chance that a 1-g subsample would have concentrations much lower (less than 1 nCi/g) than the true mean, as exemplified by the green subsample on the left.
- Look at the curve representing the 10-g subsamples (the purple subsample): Only rarely will a result from a 10-g subsample exceed 3.
- In contrast, look at the 100-g curve (red subsample). Since that curve ends around 2.5, it is very, very unlikely that any single data result would be greater than 3.
- Larger subsamples are more likely to provide data results that are close to the true mean, as evidenced by the tighter peaks the 10- and 100-g subsamples show around the true mean.
- The bottom line is that decisions that are based on a single sample result are more likely to be in error when subsample supports are small.
- As we talked about before, metals analysis typically uses around 1 gram of soil. Deciding that a few high results represent hotspots could well be decision errors due to the skewed distribution of data from small subsamples. This is why areas initially called hotspots sometimes cannot be found upon repeat sampling.
- Sampling errors operate in the other direction too. A sample from a true hotspot might give a data result biased far lower than the true value (as illustrated by the “clean” green subsample on the left) and the hotspot would be missed.

Avoiding Decision Errors



- ▶ Pay attention to QC results in the data package!
 - Suspect sampling error due to micro-scale within-sample heterogeneity when
 - Lab duplicates do not “match”
 - Matrix spikes/matrix spike duplicates do not “match”
 - Suspect sampling error due to short-scale between-sample heterogeneity when
 - Co-located samples do not “match”

Speaker Notes

- How can we avoid decision errors?
- When laboratory duplicates and/or matrix spike/matrix spike duplicates do not match and there is wide variation in results within the data set, suspect that sampling error may be occurring at the within-sample level.
- If within-sample heterogeneity has been controlled, but co-located samples do not match, the problem is likely short-scale heterogeneity.
- When sampling error has affected a data set, making decisions based on single sample results is like flipping a coin—it is a matter of chance.
- Decisions need to be based on the data set as a whole. If the data set is large and decisions are based on the mean, or the UCL on the mean, at least some of these errors could cancel out. But typical discrete data sets are much too small for that to happen.
- So work plans such as Quality Assurance Project Plans (QAPPs) need to be constructed with procedures that control for heterogeneity's effects and measure the degree of sampling error present.
- According to EPA Superfund guidance, data error must be measured for data to be definitive. (USEPA *Applicability of Superfund Data Categories to the Removal Program* OSWER 9360.4-21FS EPA 540-F-05-005 July 2006.) So QAPP reviewers should recommend that sampling error be quantified and controlled.

Avoiding Decision Errors (continued)



- ▶ Be wary of making decisions based on a single data point
 - Especially when traditional sample collection and handling is used
- ▶ Use ISM in field and lab!
- ▶ Ensure ISM work plans spell out procedures to detect and control sampling error

Speaker Notes

- How can we avoid decision errors?
- When laboratory duplicates and/or matrix spike/matrix spike duplicates do not match and there is wide variation in results within the data set, suspect that sampling error may be occurring at the within-sample level.
- If within-sample heterogeneity has been controlled, but co-located samples do not match, the problem is likely short-scale heterogeneity.
- When sampling error has affected a data set, making decisions based on single sample results is like flipping a coin—it is a matter of chance.
- Decisions need to be based on the data set as a whole. If the data set is large and decisions are based on the mean, or the UCL on the mean, at least some of these errors could cancel out. But typical discrete data sets are much too small for that to happen.
- So work plans such as Quality Assurance Project Plans (QAPPs) need to be constructed with procedures that control for heterogeneity's effects and measure the degree of sampling error present.
- According to EPA Superfund guidance, data error must be measured for data to be definitive. (USEPA *Applicability of Superfund Data Categories to the Removal Program* OSWER 9360.4-21FS EPA 540-F-05-005 July 2006.) So QAPP reviewers should recommend that sampling error be quantified and controlled.

Summary: Principles

- ▶ Inadequate management of soil heterogeneity produces highly variable data sets
- ▶ The “maximum concentration” notion is meaningless
- ▶ Chance data variability can be misinterpreted to represent the “true” condition for large soil volumes
- ▶ Misinterpreting data, especially single data points, can lead to costly decision errors
- ▶ The “nuts and bolts” of managing sampling error in the field and lab will be presented in Part 2

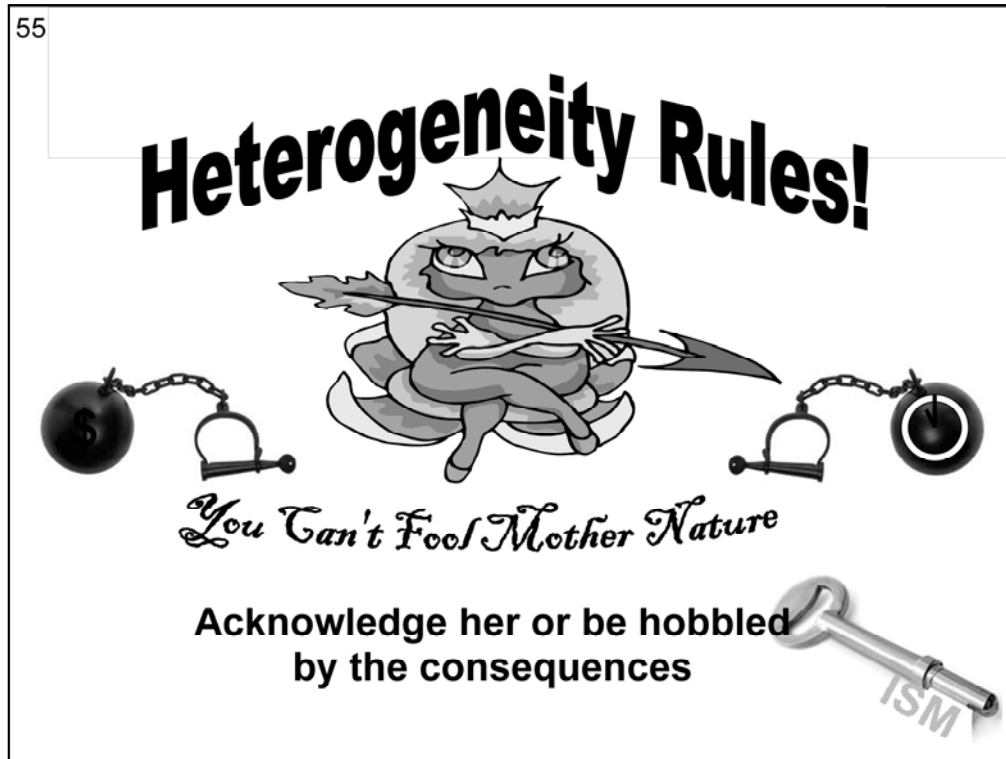
ITRC, ISM-1, Sections 5, 6, and 7

Speaker Notes

- In summary: Inadequate management of soil heterogeneity produces data sets tainted by sampling errors that manifest as data variability where data consistency would be expected.
- If single data results or very small data sets are used to make decisions, data variability and chance can produce data that might be misinterpreted as the “true” condition for large volumes of soil.
- Misinterpreting data sets can lead to costly and non-protective decision errors about risk, compliance and remediation of soils.
- Controlling sampling error is a prime feature of ISM, and more information on this will be presented throughout this webinar and in the ITRC ISM Tech-Reg document.

Supplemental Information

See ISM-1 Sections 5, 6 and 7



Speaker Notes

- As a bit of whimsy, this slide is meant to convey that heterogeneity is the natural state for soils.
- Pretending that heterogeneity doesn't exist will hobble our projects with wasted time and money.
- Incremental sampling methodology is key to managing micro-scale heterogeneity to reduce subsampling error in the lab, as well as increasing sampling densities to manage short-scale heterogeneity in the field.

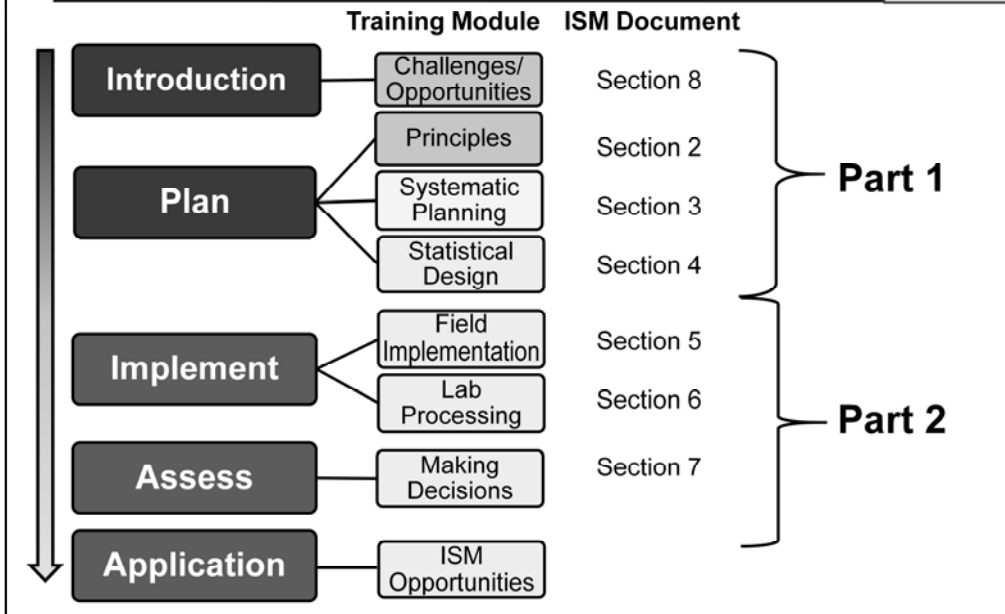
The next section of the training will delve into another aspect of ISM that must be carefully planned, which is how to select the appropriate decision unit (DU) size, configuration, and location.

Question and Answer Break



No associated notes.

ISM Document and Training Roadmap

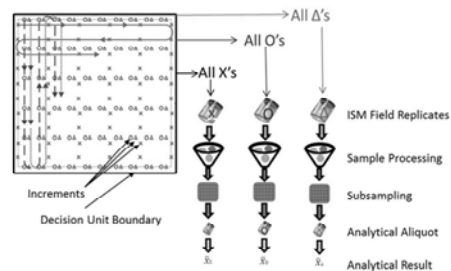


No associated notes.

Systematic Planning Learning Objectives

Learn how to:

- ▶ Conduct systematic planning steps important to ISM
 - Conceptual Site Model (CSM)
 - Risk pathways and contaminants of concern
 - Project objectives (Sampling and Data Quality Objectives (DQOs))
- ▶ Determine Decision Units (DUs)
 - Information used to develop DUs
 - Why DUs are important
 - Types of DUs
 - Real world examples (i.e., case studies)



ITRC, ISM-1, Section 3

No associated notes.

No Data Quality Objective (DQO)/Decision Units? Bad Data!



- ▶ Decision Units (DUs) – The smallest volume of soil for which a decision will be made based on ISM sampling
- ▶ Designating DUs – arguably most important aspect of ISM from a regulatory perspective
 - Selection of DUs determines
 - Where samples are being collected
 - How many
 - DU selection determines whether the data are able to satisfy the project objectives, both sampling objectives and data quality objectives

No associated notes.

Systematic Planning and Implementation

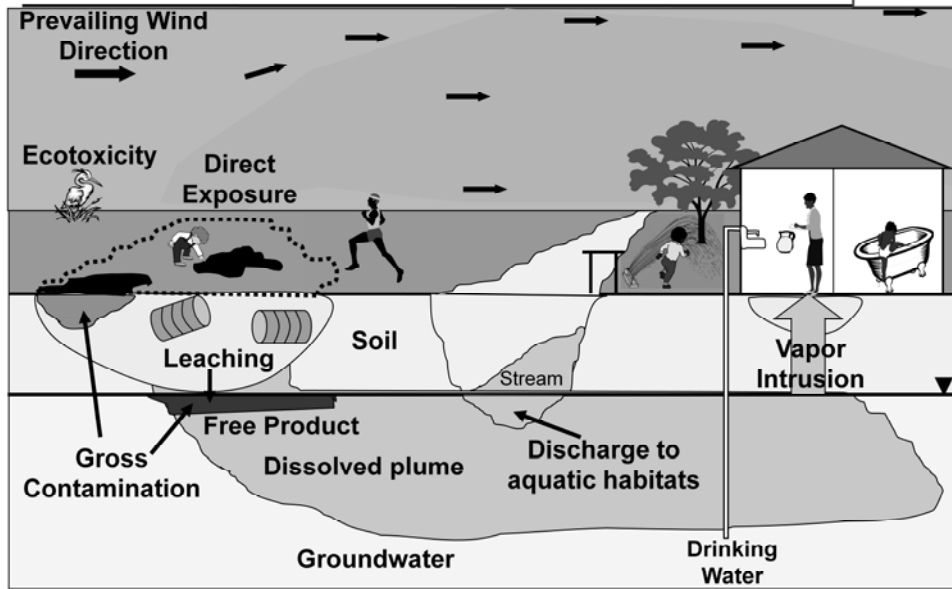


- ▶ Develop Conceptual Site Model (CSM)
- ▶ Identify contaminants and project objectives
- ▶ Identify data needed and how it will be used
- ▶ **Define Decision Units (DUs)** Key Step of ISM
- ▶ Develop decision statements
- ▶ Collect samples to characterize DUs
- ▶ Evaluate data

ITRC, ISM-1, Table 3-1

Refer to Table 3-1 in ITRC ISM-1.

Conceptual Site Model (CSM)



ITRC, ISM-1, Figure 3-2

Typical Information: Historical Site Use - Current and Future Receptors – Contaminants of Concern – Identify Potential Source Areas – Evaluate Migration Pathways – Goals of the Investigation – Geologic Conditions

ITRC ISM-1: Figure 3-2

Data/Information Needs



- ▶ What receptors and pathways are being evaluated?
- ▶ What are your sampling objectives?
- ▶ Are there multiple sampling objectives that must be met?
- ▶ What is the scale of decision making?
- ▶ What population parameter is of interest?

***The key is the volume over which
the mean should be estimated.***

No associated notes.

Example Sampling Objectives



- ▶ Estimate the mean concentration of contaminants in a pre-determined volume of soil (i.e., DU)
- ▶ Delineate the extent of contamination above screening levels
- ▶ Estimate the potential risk to receptors posed by the soil contamination
- ▶ Evaluate background metals concentrations in soil
- ▶ Confirmation sampling following remediation

No associated notes.

Designating Decision Units (DUs)

- ▶ Information used to develop DUs
- ▶ Why DUs are so important
- ▶ Types of DUs
- ▶ Examples



Stakeholder Agreement

Information that can be used to determine DUs includes: Historical site use; aerial photos for possible source areas; existing sampling data; interviews with current or former site workers; sampling objectives; and data quality objectives.

Decision Units (DUs)



The volume of soil where samples are to be collected and decisions made based on the resulting data.

Source Areas



Exposure Areas

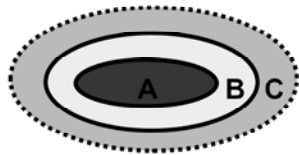


Size, shape and type of DU are an outcome of systematic planning and depend on site specific data quality objectives.

ITRC, ISM-1, Section 3.3

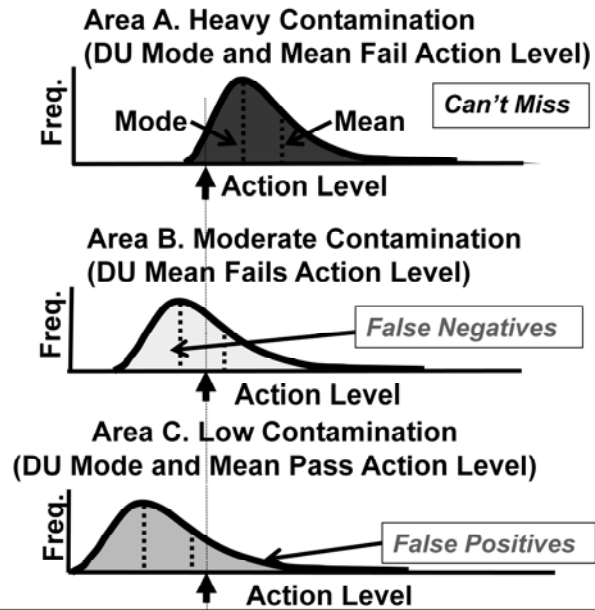
Refer to ITRC ISM-1 Section 3.3.1 Exposure Area Decision Units and ITRC ISM-1 Section 3.3.2 Source Area Decision Units

Why ISM Is Important



Example Soil Plume Map

Concentrations can vary several orders of magnitude within a DU at the scale of a discrete sample

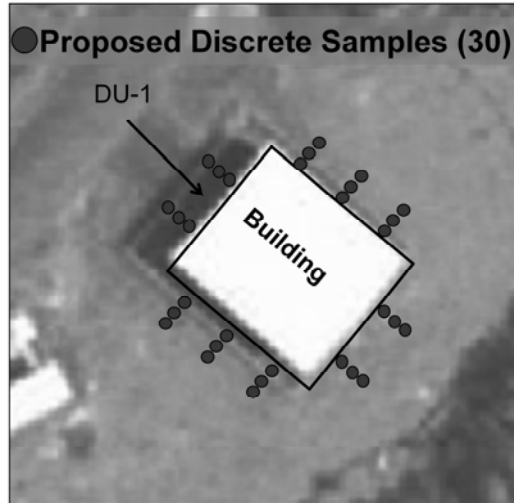


ITRC, ISM-1, Figure 2-15

Source areas can be found with discrete data, however boundaries are much more difficult. As the mean gets closer to the action level, the greater the chances discrete data will miss contamination and underestimate the mean. Take home message: Don't use discrete samples to estimate the extent of contamination!

ITRC ISM-1: Figure 2-15.

Traditional Site Investigation Approach



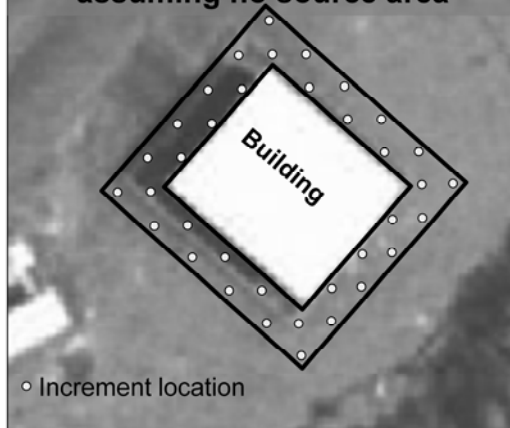
► Potential Concerns

- Inadequate number of sample points to define outward boundaries
- High risk of False Negatives and False Positives
- Confusion over single point "hot spots"
- Cost of 30 analyses
- Sample points should be randomly located for estimation of exposure point concentration (EPC)

Simple discrete design. Poor at identifying site boundaries, high risk of false negatives and false positives, potential to consider a single data point a hot spot, more expensive, and not a good design for estimating risk for an exposure area.

ISM Approach (Option 1)

Designate an exposure area DU
assuming no source area



► Advantages

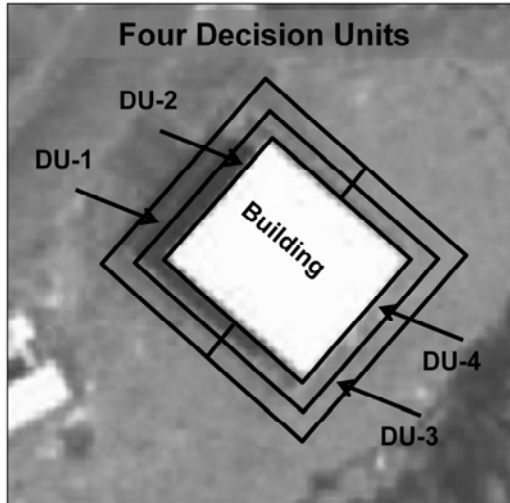
- More representative
- Risk evaluation objective identified up front
- Increments randomly and evenly spaced to minimize size of hot spot missed
- Quick and cheap if minimal contamination suspected

► Disadvantages

- Additional sampling required if DU fails

The Decision Statement: "Does contamination in soil around the perimeter of the building pose potential direct exposure concerns?" is the same as it was for the discrete design on the previous slide. This is agreed upon before the investigation. If the mean concentration is lower than a target screening level then no further action will be required. If the mean concentration exceeds a target screening level then remediation and/or further delineation will be required.

ISM Approach (Option 2)



► Advantages

- Addresses both source area and perimeter as well as directional variability if an exceedance is found
- Best approach to minimize additional sampling
- Will minimize remediation volumes if DU exceeds screening level
- If increments are collected using cores, vertical delineation is easily done with stacked DUs

No associated notes.

Suspected Lead Paint and Pesticides Around House and in Yard



Source Area DU:
perimeter of house

Exposure Area DU:
remainder of the yard

Do lead or pesticides exceed action levels around the house or in the yard?

Example Decision Units for estimating the concentration of lead from lead-based paint around a home as well as pesticides in the yard.

71

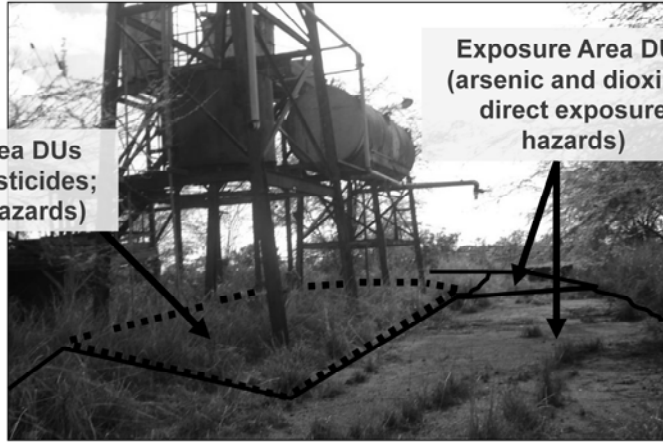
Former Pesticide Mixing Area (0.5 acres)



**Suspected heavy contamination with arsenic,
dioxins (from PCP) and leachable pesticides**

No associated notes.

Source Area and Exposure Area DU Designation



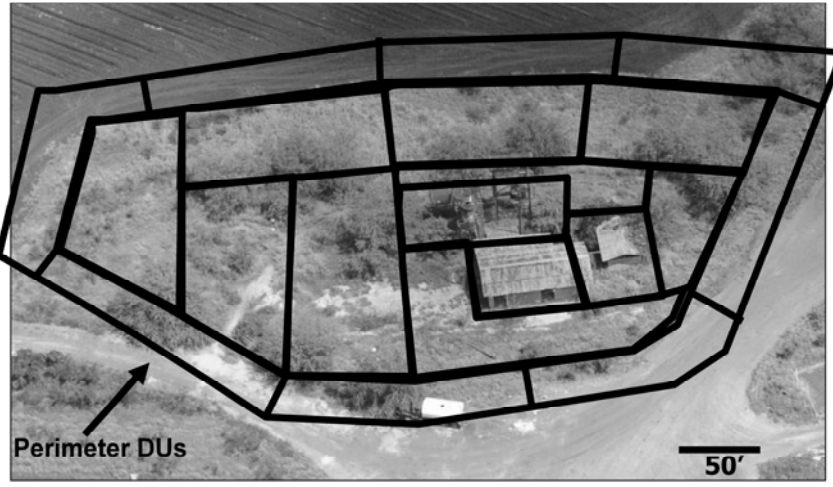
Source Area DUs
(triazine pesticides;
leaching hazards)



Exposure Area DUs
(arsenic and dioxins;
direct exposure
hazards)

Primary objective is to delineate the source area and the extent of contamination.

Use of small source area DUs and larger exposure area DUs for estimating the boundaries of contamination.

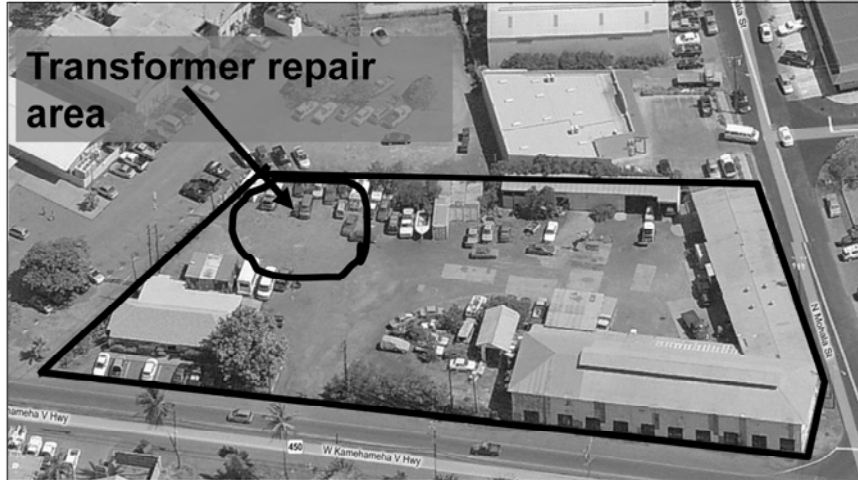
Former Pesticide Mixing Area



-  Source Area DUs: Heavy contamination + leaching
-  Exposure Area DUs: Maximum 5,000 ft²

Exposure area DUs based on a default residential lot size.
This is how the scale of decision making is tied directly to the sampling design.

Former Power Plant Proposed Community Center



Primary objective is to identify and delineate source area and extent of contamination that exceeds action levels.

100'

No associated notes.

Former Power Plant Decision Unit Designation



***Assuming 3' depth**

100'

No associated notes.

Really Big Decision Units (DU)! (400-acre former sugarcane field)



Source Area DU
(investigated separately)

Initial Screening DU

- Residual pesticide levels?
- OK for residential development?



Lot-Scale Resolution

- Hypothetical lots
- 5,000 ft² Exposure Area
- May also be required



Primary objective is to determine if property can be developed for residential use.

Refer to ITRC ISM-1 Section 3.3.7.

Really Small Decision Units???

What about the Sandbox!?



- ▶ Yard-size DUs are most often appropriate
- ▶ If acute hazards or intense exposure are being evaluated, smaller DUs may be necessary
 - Not typical
- ▶ Investigate known or suspected source areas separately
 - Remember: As sampling objectives change, so must the sampling design

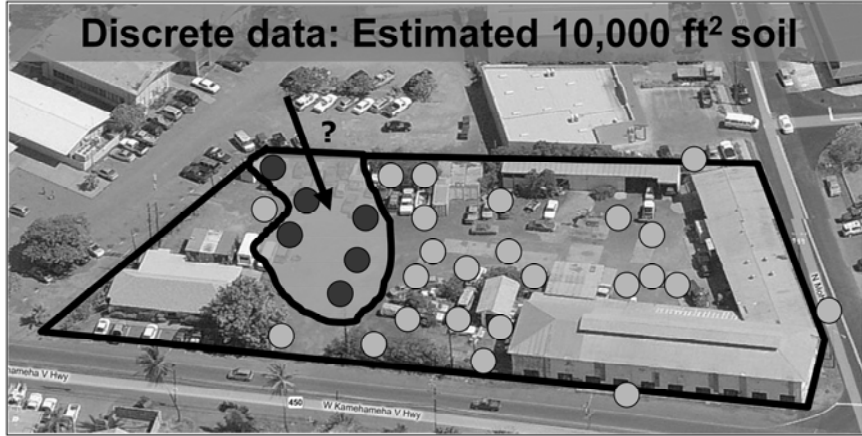
No associated notes.

78

Why DUs (and ISM) are Important (Discrete Sample Data)



● >Action Level ○ ≤Action Level



PCB sample aliquot = 30 grams (one spoonful of soil)

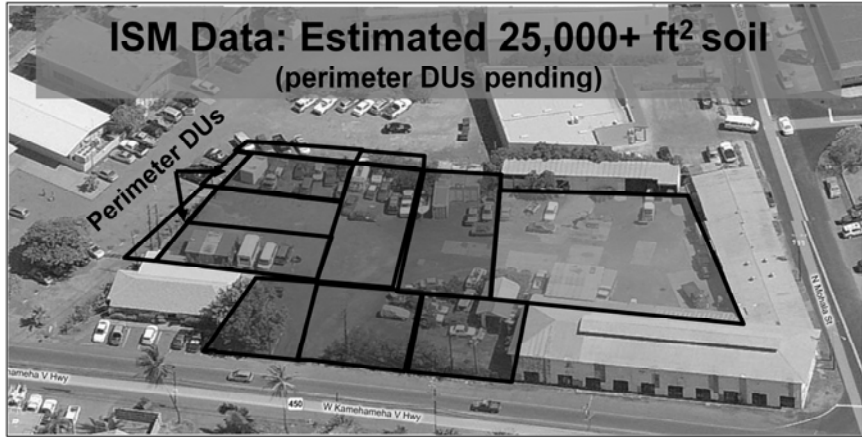
100'

No associated notes.

Why DUs (and ISM) are Important (ISM Sample Data)

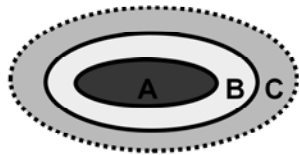


■ > Action Levels □ ≤ Action Levels



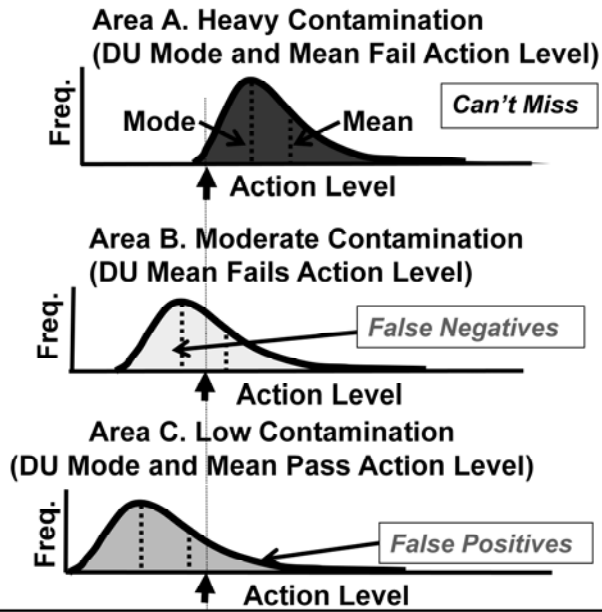
No associated notes.

Why ISM Is Important



Example Soil Plume Map

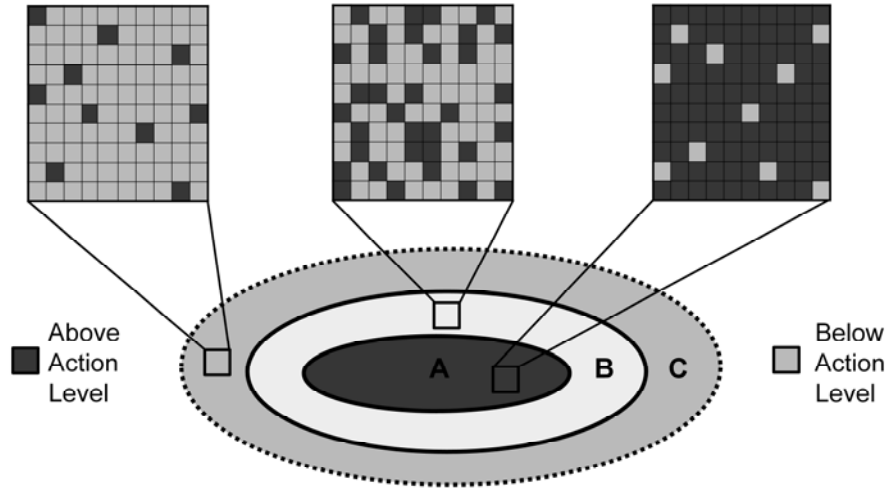
Concentrations can vary several orders of magnitude within a DU at the scale of a discrete sample



No associated notes.

Why Discrete Samples Miss Contamination in the Field

Area average PASSES (Isolated False Positives) Area average FAILS (Majority False Negatives) Area average FAILS (Isolated False Negatives)

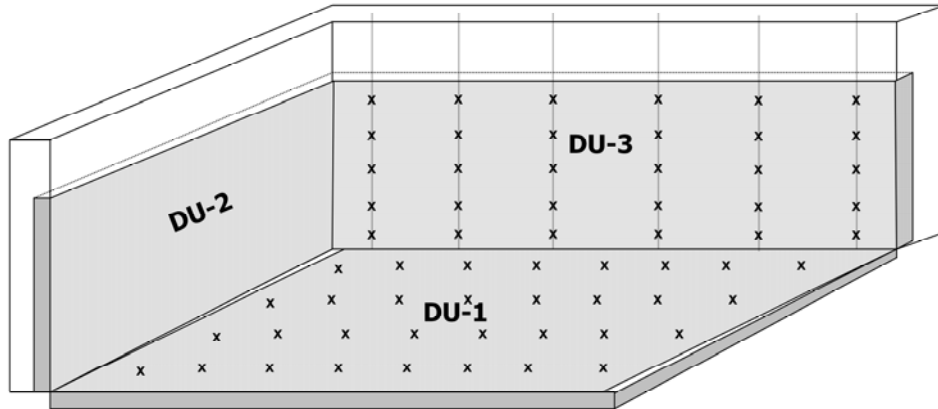


No associated notes.

Excavation Decision Units



Floor and sides tested as separate DUs



X - Increment Sampling Locations

ITRC, ISM-1, Section 3.3.6 and Figure 3-11

Refer to ITRC ISM-1 Section 3.3.6 and Figure 3-11

Stockpile Decision Units

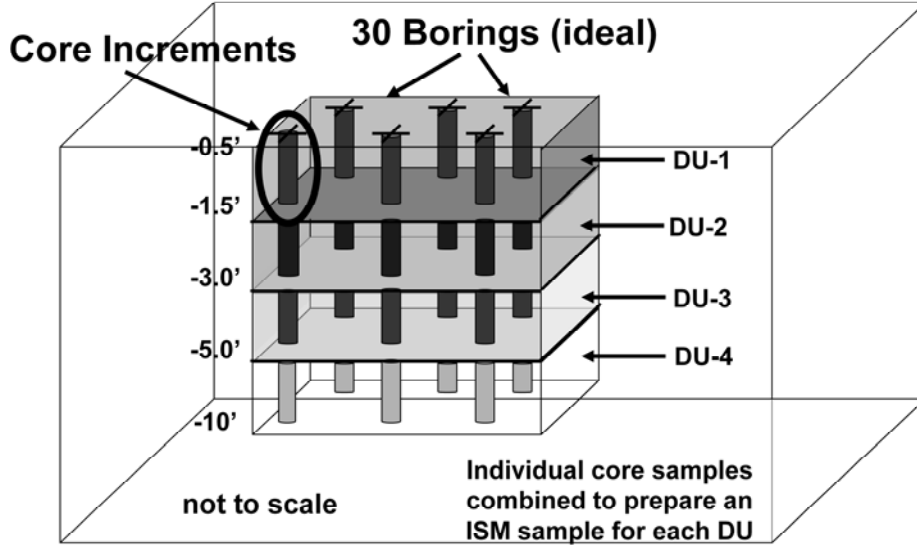


ITRC, ISM-1, Section 3.3.5 and Figure 3-10

The soil from the stockpile was to be used for fill material in a residential development, and spread out to a depth of six inches over 5,000 square foot lots. This is approximately 100 cubic yards of soil per residential lot. The stockpile was then divided into 100 cubic yard DUs.

Refer to ITRC ISM-1 Section 3.3.5 and Figure 3-10

Subsurface Decision Units



ITRC, ISM-1, Section 3.3.4 and Figure 3-8

Refer to ITRC ISM-1 Section 3.3.4 and Figure 3-8

ISM Case Study – Florida Golf Course



Decision Units designed to test various aspects of ISM, not characterize the entire golf course.

Decision Unit (DU) Highlights



- ▶ Determining DU size and location
 - Use all available information
 - Determine Data Quality Objectives
- ▶ Establish DUs with risk assessment and remedial goals in mind from the start
- ▶ Many random increments required (30 to 50+)
 - Capture the effects of heterogeneity
 - Characterize a DU

No associated notes.

Decision Unit Highlights (continued)



- ▶ ISM samples
 - More efficient and cost effective method
 - Minimize the chance of missing hot spots in the DU
 - Represent larger volumes, i.e., DUs
- ▶ Tight grids of screening data can be useful to locate suspected source areas for better DU designation, if needed

No associated notes.

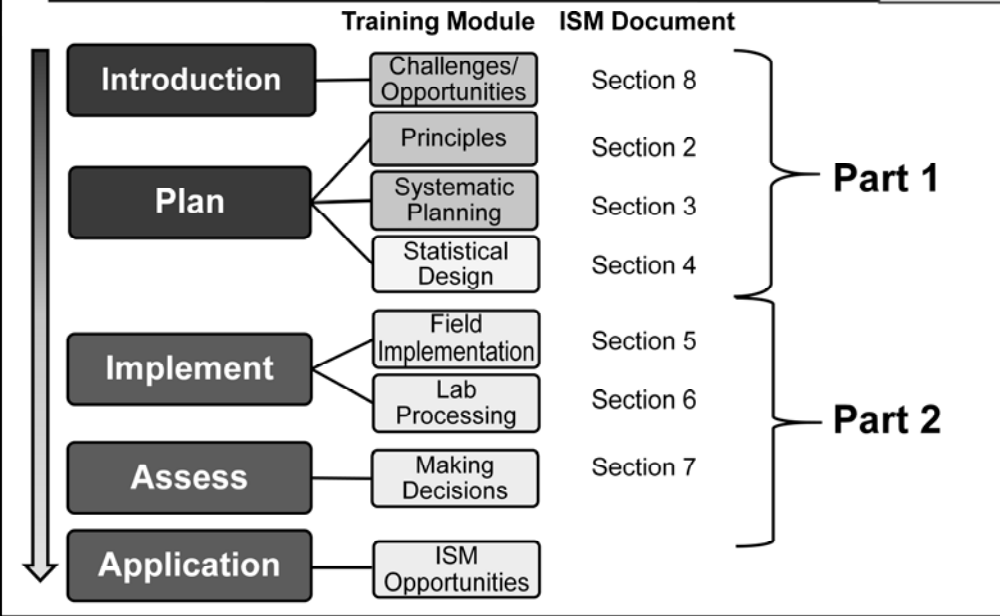
Summary: Systematic Planning



- ▶ Conduct Systematic Planning
 - It's important to develop a CSM before beginning a sampling design
 - Be sure that your sampling design will achieve your sampling objectives
 - Be certain that your sampling design will provide the kind of data necessary to fulfill the sampling objectives
- ▶ Decision Unit designation
 - Make sure that all site information has been used to develop your DUs
 - Be sure that your scale of decision making aligns with your sampling objectives

No associated notes.

ISM Document and Training Roadmap



No associated notes.

Statistical Design Learning Objectives



Learn how to

- ▶ Answer common questions about ISM related to
 - Sampling design
 - Data analysis

- ▶ Expand your understanding of
 - Statistical theory
 - Simulation studies conducted by the ITRC ISM Team

Access additional information on statistical design for ISM

ITRC ISM-1

Other training modules

Questions – Data Analysis

1. Does a single ISM sample provide a *reasonable* estimate of the mean?

Section 4.2.1



2. Can a 95UCL be calculated with ISM data?

Section 4.2.2

What is the statistical foundation for ISM?

95UCL = 95% Upper Confidence Limit of the mean

Note that section numbers of the document are given for more detailed discussions of these topics.

If the answer to #2 is yes – then the next question is - what UCL method should be used?

Questions – Sampling Design

3. What sampling design should I use?

Section 4.3.4.2

4. Is it reasonable to assume that concentrations are similar across DUs?

Section 4.4.2

5. Can background and site data be compared using ISM?

Sections 4.4.3.3 and 7.2.4

- Sampling designs include multiple features: sampling pattern, number of increments, and number of replicates.
- Extrapolation requires assumptions that the mean or variance are the same across multiple DUs.

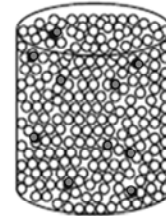
1. Does a single ISM provide a reasonable estimate of the mean?

Answer:

- It depends how much error we are willing to accept. Under some circumstances, one ISM sample can substantially underestimate the actual mean concentration.

► Why would someone collect just 1 ISM?

- UCL not required by regulator
- Save time and expense
- Assumption that more sampling wouldn't change the decision. For example
 - Variance among individual increments is low
 - Mean of DU is far above or below an action level

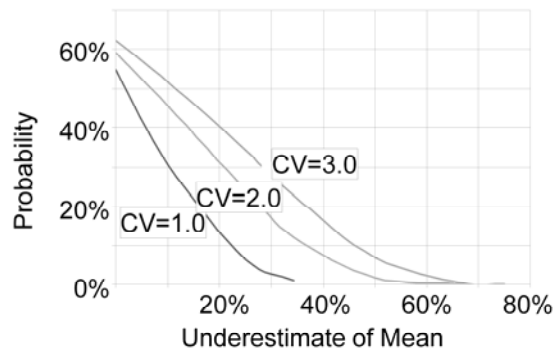


ITRC, ISM-1, Section 4.2.1

Section 4.2.1

Think of each ISM result (or “replicate”) as providing one point estimate from a distribution of possible means – the arithmetic mean of that distribution, also called the “grand mean”, is equal to the population mean. No sampling design yields a perfect estimate of the mean. The magnitude of the error in the estimate increases as: 1) the number of replicates decreases; and 2) the variance of the distribution of means increases.

1(b). How “badly” might I underestimate the mean?



CV	Frequency	Magnitude	True Mean	Estimate
1	33%	10%	400 ppm	≤ 360 ppm
2	33%	20%	400 ppm	≤ 320 ppm
3	25%	30 - 60%	400 ppm	160 - 280 ppm

***Coefficient of variation (CV) = St Dev / mean**

ITRC, ISM-1, Section 4.2.1, Figure 4-2

Figure 4-2, Section 4.2.1

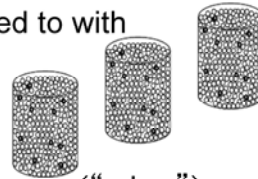
The CV in each case is based on the standard deviation of the “underlying distribution”. We generally do not know this SD (and therefore, cannot calculate the CV) since we do not measure concentrations in each increment that is composited to generate the ISM. The CV may be estimated if we also had discrete sampling.

2. Can a 95UCL be calculated?

Answer:

- Yes, even with as few as 3 ISM samples (replicates).

- ▶ Need at least 3 replicates ($r \geq 3$)
- ▶ Supported by theory and statistical simulations
- ▶ Fewer methods are available than we are used to with discrete sampling:
 - Chebyshev
 - Student's- t
- ▶ Each ISM result provides an estimate of the mean ("x-bar")
- ▶ Parameter estimates are calculated directly from ISM data



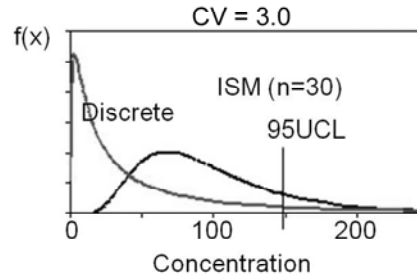
Section 4.2.2

UCL methods used with specific distributions (e.g., lognormal, gamma) would require larger sample sizes than are typically available with ISM data.

UCL methods based on bootstrap resampling would also require larger sample sizes.

2(b). How do I choose a UCL method?

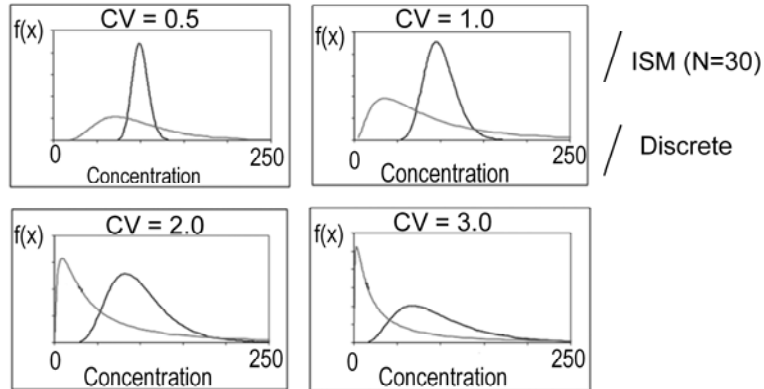
- ▶ Consider performance measures (informed by simulation study)
 - Coverage (probability $UCL > \text{mean}$)
 - Magnitude of difference between UCL and mean



- ▶ Recognize the key to performance is variability
 - Distribution of discretely \neq Distribution of ISM results
 - $CV = \text{standard deviation} / \text{mean} = 3.0$ refers to distribution of discretely
 - With only ISM results, assumptions about variability are very uncertain

In these examples, the true mean is meant to be 100. The pink distribution represents discrete data; the blue distribution represents ISM data. Chebyshev will yield a higher UCL than Student's t . This allows for better coverage but also a greater magnitude of difference between the UCL and mean. Performance varies depending on the underlying distribution's shape and variance. This makes the choice difficult because ISM data do not give much information about the underlying distribution.

Distribution of Means (ISM Replicates)



- ▶ ISM distribution variance is smaller
- ▶ ISM distribution shape becomes more non-normal with increasing CV of discrete distribution

ITRC, ISM-1, Figure 4-3

Central Limit Theorem suggests that the distribution of means approaches normality with increasing sample sizes (number of increments). However, as demonstrated here, deviations from normality are apparent for the bottom two scenarios (CV = 2 and 3). This has implications for the performance metrics of the Student's-t UCL, which provides sufficient coverage so long as the assumption of normality (of the mean) holds true.

Coverage Probabilities



UCL Method	Dispersion Among Individual Increments		
	Low (CV <1.5 or GSD <3)	Medium (1.5 < CV < 3 or 3 < GSD < 4.5)	High (CV >3 or GSD >4.5)
Student's- <i>t</i>	Yes	No	No
Chebyshev	Yes	Yes	Maybe

CV based on underlying distribution of increments

- ▶ Both methods provide desired 95% coverage when variability is low
- ▶ Chebyshev has more consistent 95% coverage for medium and high variability
- ▶ Increasing r (>3) and n (>30) provides marginal improvement in coverage for Chebyshev, but no improvement for Student's- t

ITRC, ISM-1, Table 4-4, Sections 4.3; Appendix A

This table reflects consensus results from simulation studies conducted by several members of the ISM team. This table specifically summarizes simulations of many thousands of applications of an ISM sampling protocol to hypothetical DUs. We assumed that the distribution of increments was lognormally distributed with CVs ranging from < 1.5 to > 3. ISM sampling protocols were also varied to investigate performance using different numbers of increments and replicates.

ITRC, ISM-1, Table 4-4, Sections 4.3.1.1, 4.3.3.1, 4.3.4.1; Appendix A.6.6, Fig. A-1, A-10, A-14, A-15, A-19, A-21

How much higher is Chebyshev?

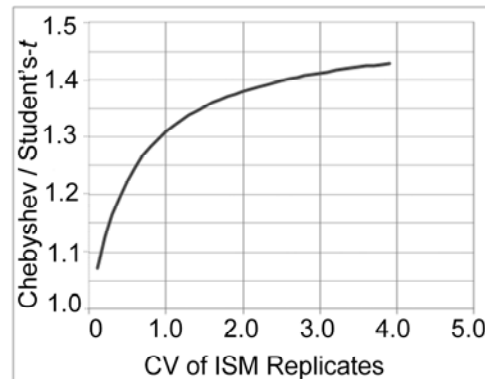
- ▶ Chebyshev will tend to yield 10-45% higher UCLs than Student's- t depending on the CV of 3 replicates
- ▶ Example: Student's- t = 100 ppm, Chebyshev = 110 -145 ppm

Chebyshev

$$UCL = \bar{X} + \left(\sqrt{\frac{1}{\alpha} - 1} \right) \frac{s_{\bar{x}}}{\sqrt{r}}$$

Student's- t

$$UCL = \bar{X} + t_{1-\alpha, r-1} \times \frac{s_{\bar{x}}}{\sqrt{r}}$$



ITRC, ISM-1, Section 4.3.1.1

When both methods are applied to the same dataset, Chebyshev will yield 10-45% higher UCLs than Student's- t . For example, if the Student's- t UCL is 100 ppm, we might expect the Chebyshev UCL to be between 110 ppm and 145 ppm. The exact difference depends on the variability in ISM results. Here we express that variability as the ratio of the SD to the mean (i.e., the CV). This CV of ISM replicates should not be confused with the CV of increments that was presented in the previous slide on coverage. Note that the Central Limit Theorem suggests that CV of replicates is ~ 5.5 times smaller than CV of increments when each replicate is comprised of $n=30$ increments. (SD replicates = SD increments / $\sqrt{30}$, and $\sqrt{30} = 5.48$). So if CV of increments = 3.0, then CV of replicates = 0.55. At this degree of variability, Chebyshev will yield about 20% higher UCL than Student's- t .

2(c). Can I use ProUCL to calculate the 95UCL?



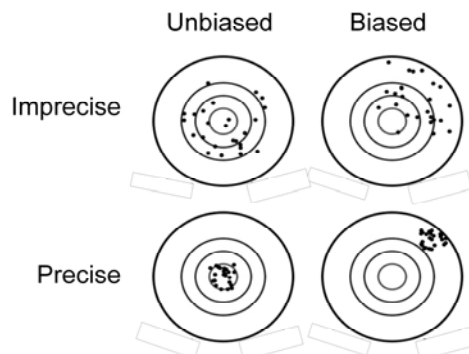
Answer:

- EPA is updating ProUCL to include an ISM module. Visit the ITRC website for additional tools.
- ▶ ProUCL was originally designed to work with discrete sample data, but is being updated to include an ISM module.
- ▶ Only Chebyshev and Student's-t UCLs are implemented for ISM datasets.
- ▶ ITRC guidance has calculator tools that work for ISM data (see ISM-1, Sections 4.2.2 at http://www.itrcweb.org/ism-1/4_2_2_UCL_Calculation_Method.html).

ITRC, ISM-1, Sections 4.2.2

ProUCL provides more options for UCL calculations, but needs $n=8$ to 10 observations in order to evaluate distributions or conduct bootstrap resampling.

Bias and Precision



- ▶ Accuracy reflects both bias and precision (reproducibility)
- ▶ These are metrics of the performance *on average*. They can only be assessed through simulation of many hypothetical sampling events – not by the results of any single ISM sampling event

ITRC, ISM-1, Section 4.3.1, Figure 4-6 and Appendix E

Figure 4-6 in Section 4.3.1 and Appendix E (Glossary)

Before answering this, it is important to review the meaning of bias and precision (reproducibility), and how they can both contribute to error.

- Bias = the tendency for a measurement to consistently over- or underestimate the actual (true) value. Together precision and bias determine accuracy.
- Precision (reproducibility) = a measure of reproducibility. Together precision and bias determine accuracy

Components of the RSD



Field

- ▶ Number of increments
- ▶ Increment collection
- ▶ Field processing
- ▶ Field splitting
- ▶ DU size and shape

Laboratory

- ▶ Lab processing
- ▶ Subsampling
- ▶ Extraction
- ▶ Digestion
- ▶ Analysis

- ▶ Simulations used to explore alternative sampling designs did not attempt to isolate sources of error
- ▶ In Day 2 of ISM training, methods will be presented for quantifying and reducing relative errors associated with field and lab practices that contribute to RSD

Error (Variability) – not “mistake”

Speaker Bullets – not comprehensive; theoretical sources of error discussed in Chapter 2 (Section 2.2) ISM controls these better than discrete

It is useful when evaluating data to consider the steps in the process where errors (variability) may have been introduced

If error is determined to be unacceptable in a given data set – one or more of these sources may be at fault

2(d). What can we infer from the RSD?

Answer:

- Without data to quantify sources of lab and field error that contribute to RSD, it is difficult to be conclusive. We expect that low RSD is an indication that steps to reduce error are successful. The 95UCL coverage does not depend on the RSD.

- ▶ RSD is the ratio of statistics calculated from ISM replicates
 - $RSD = SD / \text{mean}$
- ▶ If the goal is to make sure that the mean is not underestimated, a 95UCL should be calculated regardless of whether the RSD is high or low

ITRC, ISM-1, Section 4.3.4.4

Section 4.3.4.4

RSD measures reproducibility, not accuracy of the results.

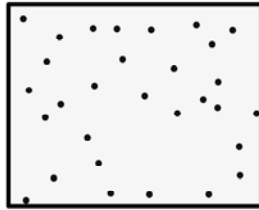
A high RSD can be caused by lab error and should be investigated. A low RSD suggests the absence of lab error, but does not necessarily indicate that the results are sufficiently accurate to avoid decision error.

3. Is there a preferred ISM sampling design?

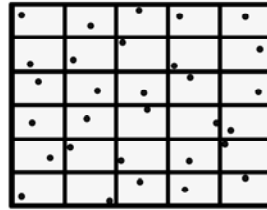


Implementation of sampling designs requires coordination between the statistician and field collection team. If the statistician provides a set of coordinates (selected at random), the field team will first place flags to match those locations in the field.

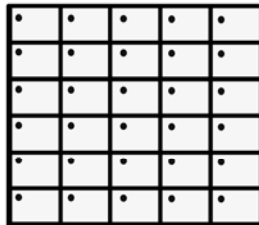
3. Is there a preferred ISM sampling design?



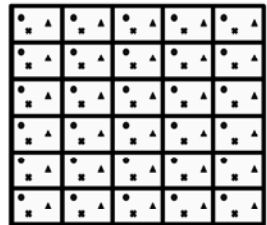
Simple Random



Random within Grid



Systematic



Systematic (3 replicates)

ITRC, ISM-1, Section 4.3.4.2

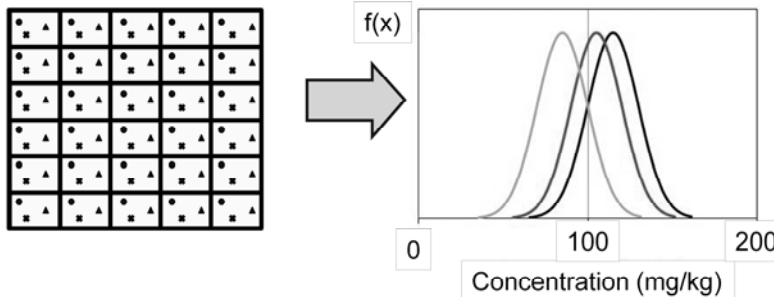
Section 4.3.4.2

With ISM, we can consider each of the 30 points to represent an individual increment. Collectively, the set of increments are combined to form a single composite sample that yields 1 ISM result. In the four sampling designs shown, each of these patterns is a form of random sampling. This means that the exact locations of the individual increments can change with each new sampling event. With #1, even though the increments are equidistant, the location of the first increment is selected a random from within the grid cell. With #2, three sampling events are shown (circle, square, triangle), each with a different random “starting” location. It is clear how the density of the samples (or “spatial coverage”) can be very high with ISM designs that involve multiple replicates.

3. Is there a preferred ISM sampling design (continued)?

Answer:

- Each random sampling design yields unbiased estimates of the mean and is an acceptable approach in most situations.
- Systematic random sampling is most often used because it is the easiest to implement random sampling,



ITRC, ISM-1, Section 4.3.4.2

Section 4.3.4.2 (sampling design) and Section 4.3.1.2 (bias)

Random sampling from a population generates an estimate of the mean concentration of that population with desirable statistical properties.

Bias is one metric used to evaluate the performance of a parameter estimation method. A sampling method is unbiased if, when repeated many times, the parameter estimate equals the population parameter. This examples shows the distribution of increments for three sampling events applied to a DU with a true (population) mean of 100 mg/kg. For purposes of presentation, we also assume that the distribution is normal. While no individual ISM is centered on the mean, on average (if repeated many times), we would expect the estimate of the mean to equal 100 mg/kg.

3(b). How many increments?

Answer:

- n = 30: generally, 30 increments per ISM sample provide good results. Lower numbers are discouraged and higher numbers provide diminishing improvement in statistics.

- ▶ As the number of increments increases:
 - spatial coverage improves (greater sample density)
 - lower variability in ISM results (smaller standard deviation)
 - 95UCL will tend to be closer to the mean
- ▶ Size of DU can be a consideration – large DUs may require more increments

10 20 **30** 40 50 60 70 80 90 100

ITRC, ISM-1, Sections 4.3.4.1 and 5.3.1

Section 4.3.4.1

Sample support and spatial coverage are usually optimized at about 30 increments. Lesser numbers compromise the ability of ISM to address sampling error; larger numbers provide limited improvement in error reduction, though a “denser” sampling grid. In the simulations, we calculated the difference between the UCL and mean using the “relative percent difference”, $RPD = (UCL - \text{mean}) / \text{mean}$

The simulations that support recommendations given in the ISM guidance apply to DUs of any size because heterogeneity was represented by statistical distributions that did not depend on the size of the DU. Some practitioners of ISM have found that increasing the number of increments to more than n=30 for larger DUs (e.g., 1 acre or more) can provide greater confidence that, collectively, the set of increments collected in the DU will represent subareas of high and low concentrations in the appropriate proportions. As discussed in Section 5.3.1 of the ISM guidance, as the DU gets significantly larger, the amount of distributional heterogeneity may increase. In these cases, depending on site specific knowledge, the conceptual site model (CSM), and data quality objectives (DQOs), it may be necessary to increase the number of increments per DU to n=50 or more. In practice, collection of a greater number of increments in each DU typically reduces the variation among replicate samples. Alternatively, splitting larger DUs into two or more smaller DUs should be considered. It is not normally necessary to increase the number of increments unless there is reason to believe the DU has more distributional heterogeneity than can be controlled with 30–50 increments.

3(c). How many replicates?

Answer:

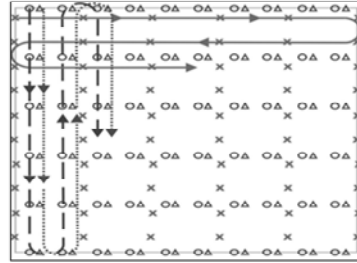
- $r=3$: for most DUs, three replicates is sufficient.

- ▶ Minimum number to calculate standard deviation (and 95UCL) of ISM results
- ▶ More replicates will produce a 95UCL closer to the actual mean, but may not be cost-effective unless the result is near the action level



Which Would You Choose and Why?

- A. $n = 30, r = 1$
- B. $n = 90, r = 1$
- C. $n = 30, r = 3$ (so $30 \times 3 = 90$)



Scenario	Spatial Coverage	Analysis Cost	Estimate of Mean	Estimate of Variance
A	Low	Low	Yes	No
B	3 x A	A	Yes	No
C	3 x A	3 x A	Yes	Yes

With ISM, selecting the sample sizes requires weighing advantages and disadvantages of factors such as: spatial coverage of the DU, costs (including field and analytical), and whether or not an estimate of the variance can be obtained to calculate a 95UCL. The table above compares and contrasts three scenarios. Green shading is more favorable, orange shading is less favorable. In general, we would opt for a sampling design that provides an estimate of the variance.

4. Can I extrapolate results across DUs?

- ▶ Unsampled DU – extrapolate estimate of mean
- ▶ DU with 1 ISM – extrapolate estimate of variability
 - Standard deviation (SD)
 - Coefficient of variation (CV)



DU-1 $\stackrel{?}{=}$ DU-2



ITRC, ISM-1, Section 4.4.2

Section 4.4.2

There are different assumptions associated with each, and therefore different answers...

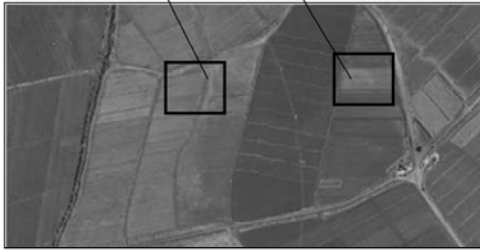
If you extrapolate variability, generally use the CV instead of the SD. This is because the CV estimates standard deviation based on the estimate of the mean: $CV = SD/\text{mean}$. This is more consistent with contamination that has a non-normal (or “positively skewed”) distribution (e.g., lognormal, gamma). Extrapolation of the SD is only recommended if the distribution of the concentrations of individual increments is approximately normal.

4(b). Extrapolation of the Mean

Answer:

- You are assuming that the mean concentration in the unsampled DU(s) is the same as in the sampled DU.

DU-1 ? DU-2



- ▶ DU-1:
 - Mean = 100
 - SD = ?
- ▶ DU-2:
 - Extrapolation:
Mean = 100



ITRC, ISM-1, Section 4.4.2

Extrapolation is usually reserved for very large sites where sampling all DUs is prohibitively expensive.

There is usually no confirmation of this assumption, so the rationale for the assumption must be very strong.

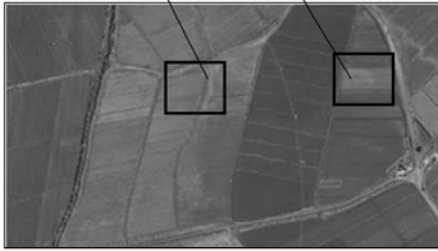
Note that there is nothing magic about ISM that diminishes the uncertainty of extrapolating from sampled to unsampled areas. If you wouldn't do it with discrete data, you shouldn't do it with ISM.

4(c). Extrapolation of the Variance

Answer:

- You are assuming that the heterogeneity in contaminant concentrations is similar in all of the DUs.

DU-1 ? DU-2



ITRC, ISM-1, Section 4.4.2

▶ DU-1:

- Mean = 100
- SD = 50
- CV = SD/mean = 50/100 = 0.5



▶ DU-2:

- Mean = 400
- Extrapolation:
CV = 0.5 = x / 400
therefore
SD = x = 200

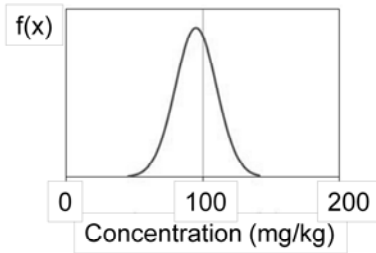
Extrapolation of variance is valid only if the distributions of concentrations in each DU are statistically similar.

If you extrapolate variability, generally use the CV instead of the SD. This is because the CV estimates standard deviation based on the estimate of the mean: $CV = SD/mean$. This is more consistent with contamination that has a non-normal (or “positively skewed”) distribution (e.g., lognormal, gamma). Extrapolation of the SD is only recommended if the distribution of the concentrations of individual increments is approximately normal.

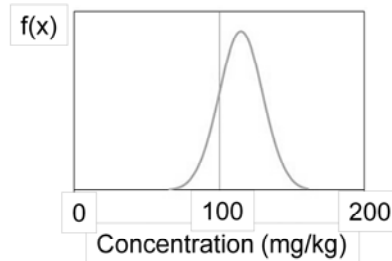
5. Can background and site ISM data be compared?

Answer:

- Yes, but statistical tools for comparison are limited.



Background



DU-1

- ▶ Each data sets consists of ISM samples, preferably generated with similar sampling designs

ITRC, ISM-1, Section 4.4.3.3

Section 4.4.3.3

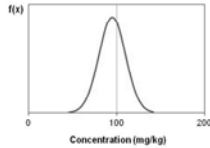
The concept is to compare distributions. With ISM, we will generally have too few data points (e.g., $r=3$) to determine distribution shapes, or to use non-parametric methods. One cannot compare discrete data to ISM data because they represent the distribution of concentrations very differently.

5. Can background and site ISM data be compared?

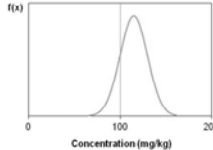
Answer:

- Yes, but statistical tools for comparison are limited.

Background



DU-1



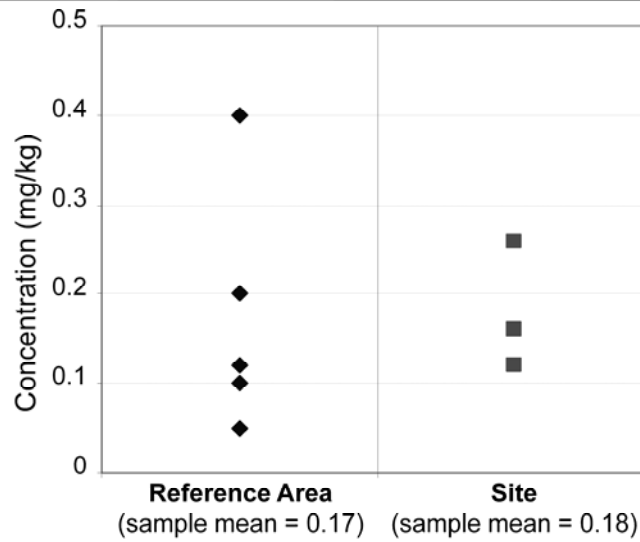
- ▶ Equal central tendency (mean, median) ?
- ▶ Equal upper tails ?
- ▶ Hypothesis testing is limited to parametric tests of the mean:
 - Assume distribution shape
 - Use estimates of mean, SD, and number of replicates
- ▶ Cannot test upper tails with ISM data

ITRC, ISM-1, Section 4.4.3.3

The concept is to compare distributions of ISM results. With ISM, we will generally have too few data points (e.g., $r=3$) to determine distribution shapes.

Because of the small sample size, statistical power is too low to allow for non-parametric hypothesis tests like Wilcoxon Rank Sum (Mann Whitney).

5. Example Background Comparison



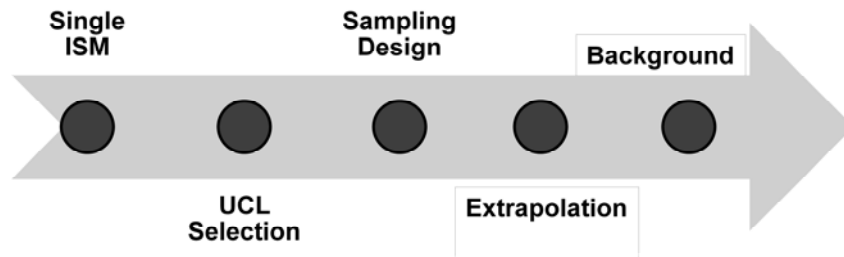
ITRC, ISM-1, Section 7.2.4, Figure 7-1

Figure 7-1, Section 7.2.4

Dot plots can be a useful graphic for presenting ISM results. While not the same as a formal statistical test, the graphic presents information in a manner that can support decisions.

Recap of Learning Objectives

See: Section 4 and Appendix A



This Module 4 brings the Day-1 training to an end. The focus of this module has been on specific questions that can be addressed through statistical analysis of ISM data. For more detailed discussion of these concepts and an overview of the simulations that were performed to support the recommendations that have been provided, please refer to Section 4 and Appendix A of the document.

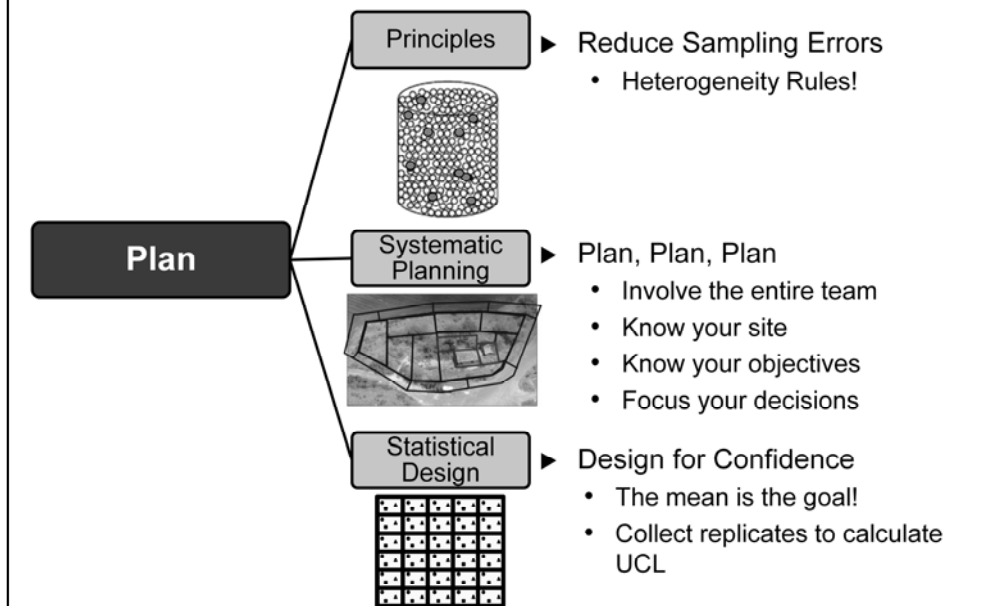
Summary: Statistical Design



- ▶ Mean or 95UCL from ISM data may be used to make decisions about a site
- ▶ 3 replicate samples provide adequate information to calculate a 95UCL
- ▶ Systematic random sampling is most commonly used
- ▶ About 30 increments per ISM sample is usually sufficient
- ▶ Extrapolation of the mean or variance can be very uncertain
- ▶ Comparisons between ISM data (e.g., site vs. background) are possible, with caution

No associated notes.

118 **ISM Part 1 – Summary
Principles, Systematic Planning, and
Statistical Design**



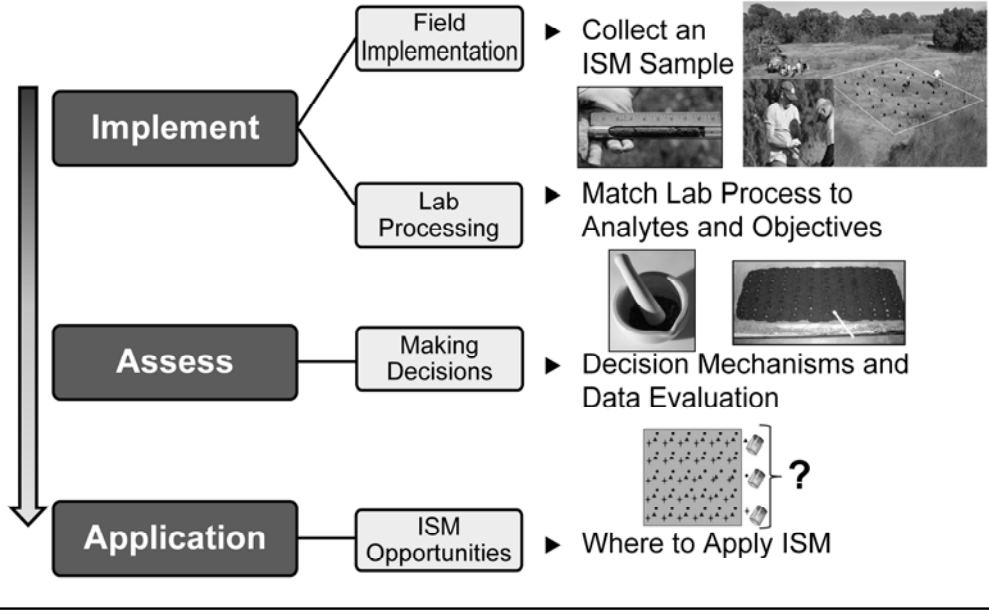
Today's training included Modules on:

Soil and Principles – Be aware of issues related to heterogeneity and sampling errors

Systematic Planning - Involve the entire team, regulators, consultants, responsible parties in critical elements (e.g. conceptual site model, establish sampling objectives and decision units.) Sampling objectives should drive your sampling design, and the scale of decision making should align with sampling objectives.

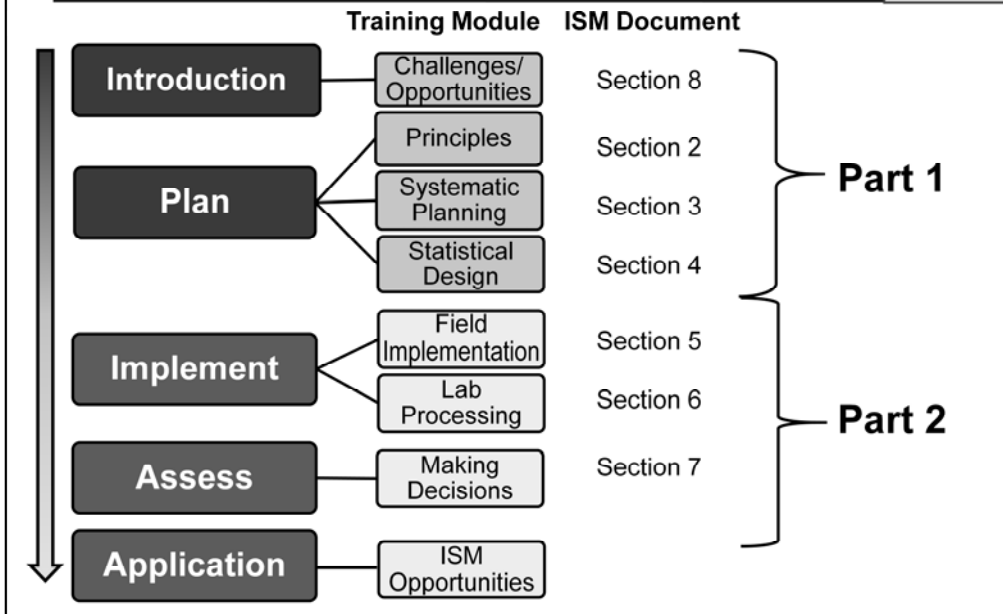
Statistical Design - Provides the statistical foundation and describes why ISM provides a reasonable mean, describes a good ISM sampling design, and informs you how ISM provides 95% UCL.

ISM Part 2 Preview Implement, Assess, and Apply



No associated notes.

ISM Part 1 Summary and Part 2 Preview



Part 2 Module includes:

- Collecting Field Samples
- Laboratory Processing & Analytical Issues
- Using and Applying ISM Data

Thank You for Participating



- ▶ 2nd question and answer break
- ▶ Links to additional resources
 - <http://www.clu-in.org/conf/itrc/ISM/resource.cfm>
- ▶ Feedback form – *please* complete
 - <http://www.clu-in.org/conf/itrc/ISM/feedback.cfm>

Need confirmation of your participation today?

Fill out the feedback form and check box for confirmation email.

Links to additional resources:

<http://www.clu-in.org/conf/itrc/ISM/resource.cfm>

Your feedback is important – please fill out the form at:

<http://www.clu-in.org/conf/itrc/ISM/feedback.cfm>

The benefits that ITRC offers to state regulators and technology developers, vendors, and consultants include:

- ✓ Helping regulators build their knowledge base and raise their confidence about new environmental technologies
- ✓ Helping regulators save time and money when evaluating environmental technologies
- ✓ Guiding technology developers in the collection of performance data to satisfy the requirements of multiple states
- ✓ Helping technology vendors avoid the time and expense of conducting duplicative and costly demonstrations
- ✓ Providing a reliable network among members of the environmental community to focus on innovative environmental technologies

How you can get involved with ITRC:

- ✓ Join an ITRC Team – with just 10% of your time you can have a positive impact on the regulatory process and acceptance of innovative technologies and approaches
- ✓ Sponsor ITRC's technical team and other activities
- ✓ Use ITRC products and attend training courses
- ✓ Submit proposals for new technical teams and projects