Module 7

Incremental-Composite Sampling Designs for Surface Soil Analyses

What's in a Name?

MIS

Composite sampling

ISM

ICS

Incremental sampling

“Composite” or “Multi-Increment” or “Incremental” Sampling

- Composite sampling - term used since 1985 in USEPA guidance to convey idea of “pooling” for several purposes
- Multi-Increment Sampling ® – term coined by Chuck Ramsey which he has trademarked as “MIS”
- Incremental sampling methodology (ISM) – term used by ITRC due to copyright infringement concern over “MIS”; goal is to obtain average concentration over DU
- Incremental-composite sampling – USEPA term to combine ITRC’s ISM with features from USEPA’s existing compositing guidance; goals include more than the average

All Share Common Characteristics

- All refer to collecting soil samples:
  - From physically separate locations (“increments”)
  - Then pooling to form one homogenized sample (termed a “composite sample” or “incremental sample”)
- Primary difference lies in their purpose; e.g., finding an average (incremental) vs searching for hot spots (a compositing technique)
- MIS ® involves a specific incremental sampling protocol that has been optimized for explosives residues, but can be generalized to other analytes

Every soil sampling program can be improved by using some level of composite sampling in the field and/or in lab subsampling. “Improved” means increased performance for the same cost, or cost savings for the same performance.

A Long History of EPA Composite Sampling Guidance

- 6 documents since 1985 go into some depth
  - 1985 PCB Technical Guidance
  - 1995 EPA Observational Economy Series
  - 1996 Soil Screening Guidance, Part 4
  - 2002 EPA RCRA Sampling Technical Guidance
  - 2002 EPA QA/G-5S
  - 2006 EPA SW-846 Method 8330B (App. A)
- They do not cover all potential issues or details, but do provide a framework

Incremental Averaging vs. Composite Searching: Basic Differences (1)

- Goal of incremental averaging: estimate the average concentration over some defined area/volume of soil
- A defined soil area/volume is the subject of a decision of e.g., risk/no risk or exceedance/no exceedance
  - The defined area/volume is called a decision unit (DU)
  - Having 3 or more independent replicates allows calculation of an upper confidence limit (UCL) on the mean (i.e., a conservative estimate of the mean concentration)

Incremental Averaging vs. Composite Searching: Basic Differences (2)

- Common goal of composite searching: gain information about contaminants' spatial distribution
  - Accurate or conservative estimate of the concentration mean within the given area not required
  - The "given area" is called a sampling unit (SU); typically smaller than the DU
  - Composite replicates not typically used
  - Improves representativeness of "single" sample representing a small area by suppressing short-scale heterogeneity
- Composite searching typically uses fewer increments per sample than incremental averaging

“Incremental Sampling” per ITRC

A structured sampling and subsampling protocol for representative & reproducible sampling of a well-defined soil area or volume (a decision unit) to obtain a concentration value representative of the average concentration over the entire decision unit

ITRC ISM-1 document:
http://www.itrcweb.org/teampublic_ISM.asp

Other Recent Guidance Documents

- Hawaii state guidance for MIS at http://www.hawaiidoh.org/tgm-pdfs/HTGM%20Section%2004-02.pdf
- Alaska state guidance for MIS at http://dec.alaska.gov/spar/csp/guidance/multi_increment.pdf
- Army Corps of Engineers http://www.hnd.usace.army.mil/oew/policy/IntGuidRegs/IGD%209-02v2.pdf

Key Components of Any Good Sampling Design

- **Project Planning & Field Sample Collection**
  - Well-formulated & explicitly stated sampling objectives to define the decision unit (DU).
  - Determine whether data goal is only to determine DU average, or also to preserve spatial information.
- **Sample Processing and Subsampling**
  - Sample processing may begin in the field and finish in the lab, or all be done in the lab.
  - Goal: maintain chain of sample representativeness for each step thru subsampling.
  - As important for discrete designs as for ICS!

2 Key Assumptions of Incremental Averaging

(1) There is a specified volume of soil called the DU

(2) There is a decision that rests on knowing THE concentration of that volume of soil

<u>If it were possible</u>, how would THE true concentration of the DU be determined? What is done instead?

Average Concentration Estimation

- Determining average concentrations over the DU area is often a significant goal:
  1) Developing exposure point concentrations (EPC) for risk assessment purposes
  2) Demonstrating compliance with
     - Area-averaged cleanup goals
     - “Hot spot” criteria
- Fundamental concept:
  – ***How is the decision unit defined???***

Discrete Sampling Approach

- Collect one or more samples from an area, analyze via off-site laboratory, and use the mathematical average (plus an uncertainty buffer = UCL) to estimate true average
- Heterogeneity introduces uncertainty into how good that estimate is
- More samples create a better estimate
- Statistics (e.g., VSP) can be used to estimate required sample numbers **before sampling**
- Statistics (e.g., ProUCL) can be used to determine how good an estimate is **after sampling**

Issues with Discrete Sampling Approach

- How many samples are enough?
  - Won't know until after samples are collected
- Statistics (e.g., VSP): more samples than budget
- The statistical approach might assume underlying sample distribution is normal (bell-curved)
  - Assumption affects sample numbers
- Expensive
- Tendency to make decisions based on limited sample results—DANGEROUS

Incremental Averaging

- Used to cost-effectively suppress short-scale (collocated) heterogeneity
  - Estimates of the mean less uncertain & closer to true
- Multiple increments contribute to the composite that is analyzed
- Increments systematically distributed over an area equivalent to, or less than, decision requirements
- ***Effective when the cost of analysis is significantly greater than cost of sample acquisition***
- **Benefit dependent on sample processing!!**

How Does Incremental Averaging Work?

- Physical equivalent of averaging individual sample results
- Tends to “normalize” underlying distribution, allowing simpler statistics to be used
  - Student’s t tests, Student’s t UCL
- A set of composite sample results show less variability than discrete sample counterparts
- Theoretically, the more increments per sample per DU, the lower the variability in sample results
  - Sample processing is a critical factor!!

What About Dilution Concerns?

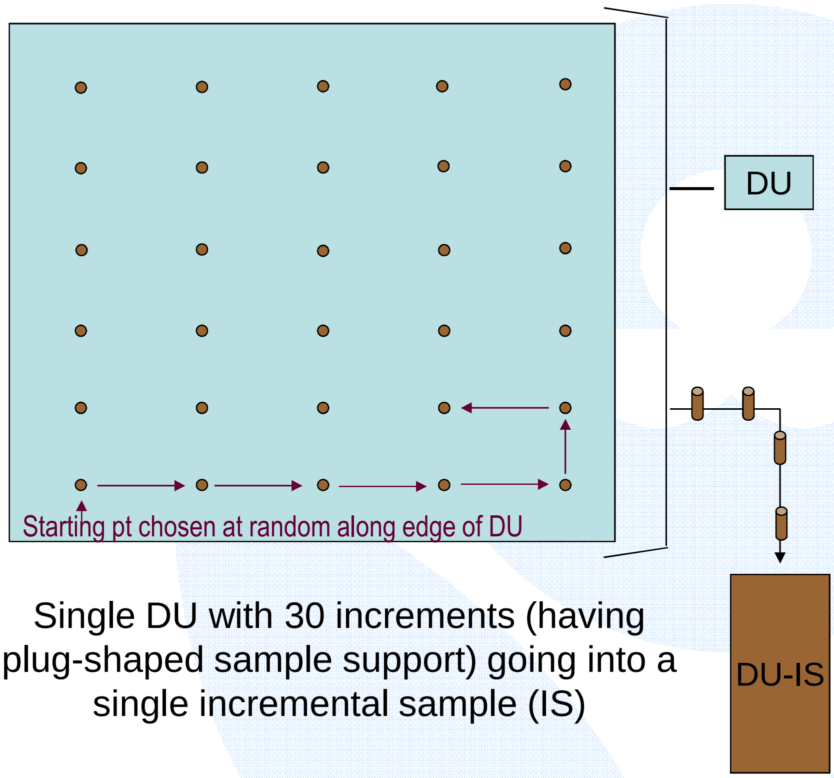
- For area-averaging goals, the concern doesn't apply
  - Goal is to get estimate of average concentration over the DU (i.e., an exposure unit)
  - Pooling increments with proper processing is physical equivalent of mathematical averaging
  - High increment density incorporates high & low concentration areas in actual field proportions
- For hot spot identification, compositing works *against* missing hot spots:
  - Hot spots also an average concept, but over smaller area
  - Compositing actually *increases* likelihood that hot spots will be incorporated into the ICS sample, raising its concentration
  - Higher sample concentration flags area for more investigation

How Many Increments Total Per DU?

- Can vary depending on size of the area & sampling design
- For incremental samples (ISs) that cover a DU with 1 sample
  - ITRC recommends ~30 as a default for general contaminants and DUs about residential size; statistical simulations also support 30
  - Large areas may need more increments to achieve sufficient density
  - More needed for areas with high spatial heterogeneity: Military energetics & metals on firing ranges need 50-100
  - Generally fewer needed to confirm low or high concentrations far from action level
  - More needed where suspect close to action level
- If 3 replicate ISs per DU used (gives a UCL), density is 90 increments per about ¼-½ acre—generally sufficient

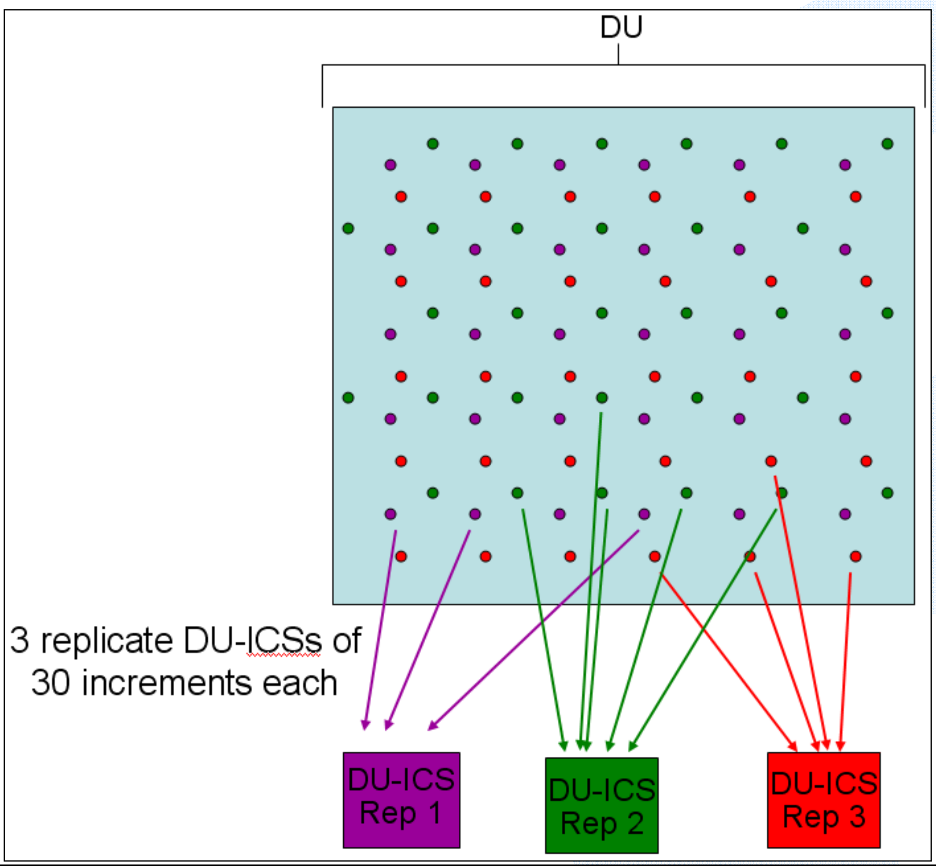
Incremental Sampling

Taking a single incremental sample over a decision unit (DU)



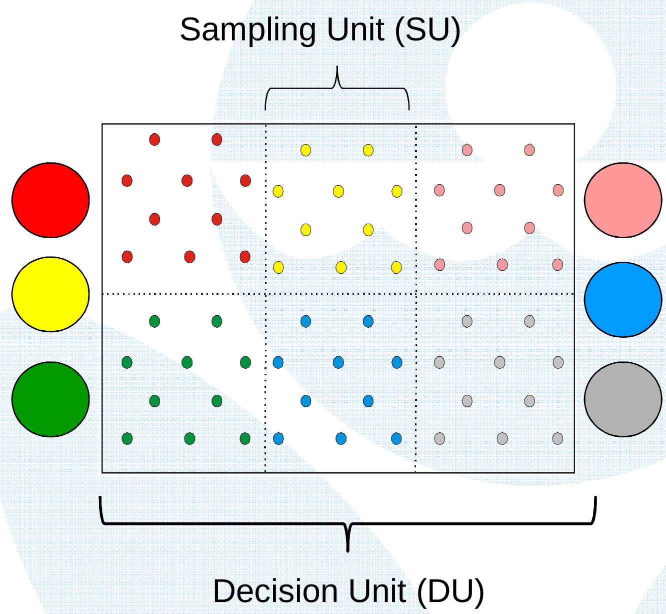
Single DU with 30 increments (having plug-shaped sample support) going into a single incremental sample (IS)

Replicate Incremental Samples



Example of A Non-Overlapping Incremental Averaging Design

- DU contains 6 SUs
- 6 composite SU samples composed of 10 increments each: 60 total
- All SU samples immediately analyzed
- Increments distributed systematically over adjacent SUs
- Preserves spatial info
- Can calculate UCL



A Non-Overlapping Incremental Averaging Design (cont.)

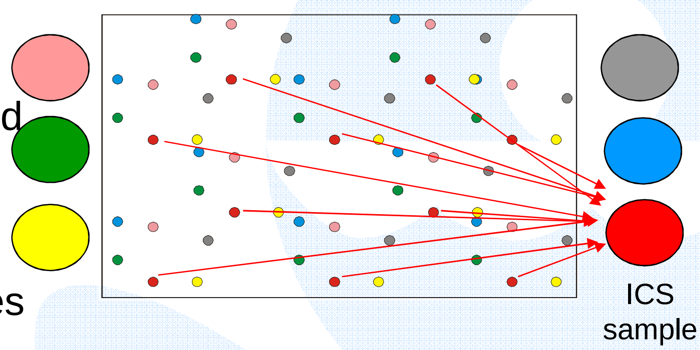
- If composites are non-overlapping, similar to discrete samples, but more representative of “area of inference”
- Increased sample support of the composite reduces data variability (reduces the skewed character of data distribution)
- These are not replicates, so data variability > variability between ISM replicates
  - Because spatial variability exists, cannot assume data will be normal!
- If have 8 or more, can calculate UCL with ProUCL
- If <8, safest to use the nonparametric Chebyshev formula

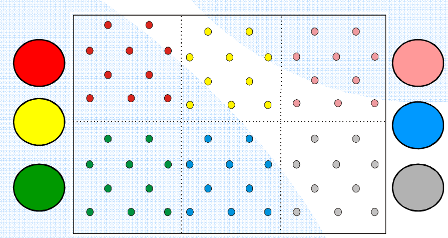
$$UCL = \bar{x} + \left(\sqrt{\frac{1}{\alpha} - 1} \right) \frac{s}{\sqrt{n}}$$

Alternate Incremental Averaging Design: Replicate DU Composites

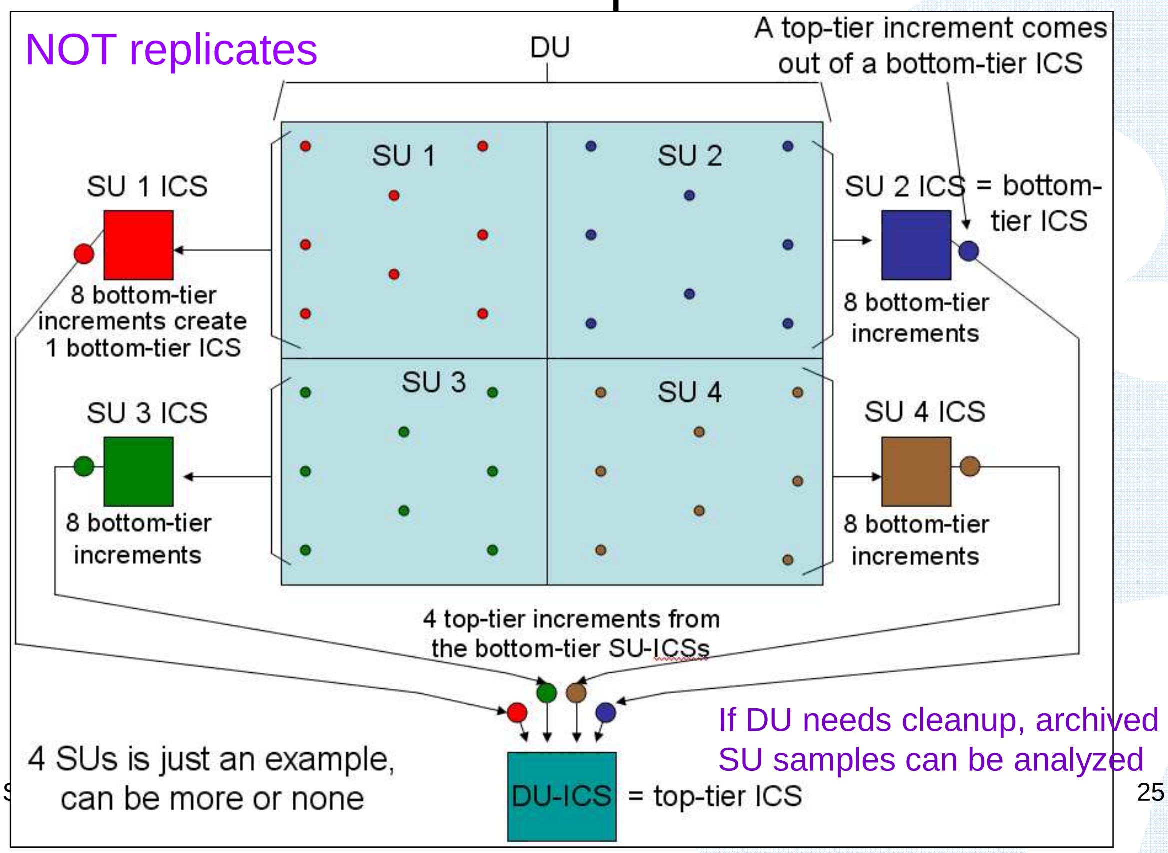
6 replicate composite samples composed of 10 increments each

- Each set of increments distributed over entire DU
  - Increments can be placed randomly or systematic
  - VSP can estimate increments & composite numbers & place samples
- Contrast against previous non-overlapping composite design
- For uniform DUs, UCL may be lower than 30-inc ISs w/ n=3 because of higher n





Alternate Incremental Averaging Design: Tiered Composites



25

Composite Searching: Looking for Problems

- Examples:
  - Does contamination exist at a site?
  - Does contamination exist at a site above levels of potential concern?
  - Are there **"hot spots"** that need to be addressed?
- Again – definitions are fundamental:
  - ***How is the decision unit defined???***
  - ***How is a "hot spot" defined???***
    - Need conc & the area over which that conc applies

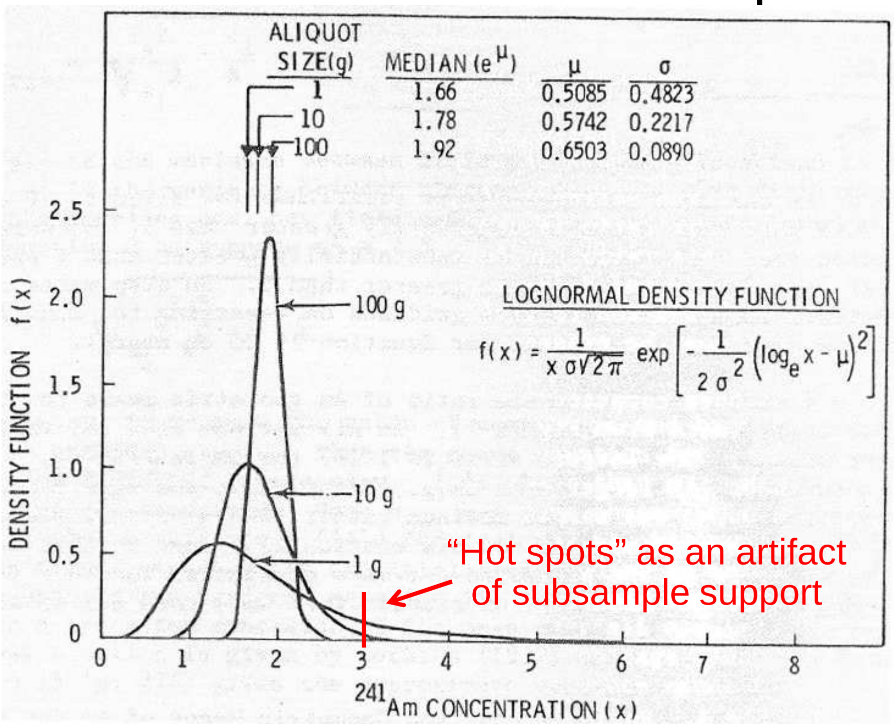
Goals Are Different From Those for Incremental-Averaging

- Not really interested in accurately knowing the average concentration
- Only interested in knowing whether concentrations are above or below some threshold
- The issue is how reliably can we identify situations when contamination is truly above the relevant threshold given our sampling strategy

Decision Units and Searching for Contamination

- When looking for evidence of contamination (e.g. SI or RI phase), decision units are often not well-defined
- For contaminants that are not naturally occurring, contamination evidence is a positive detection or positive detection above some threshold
- For contaminants that are naturally occurring, comparison is often to a background threshold value (BTV) and/or to some screening level

Recall: variability & statistical distribution is a function of soil sample mass



Adapted from DOE (1978) americium-241 study

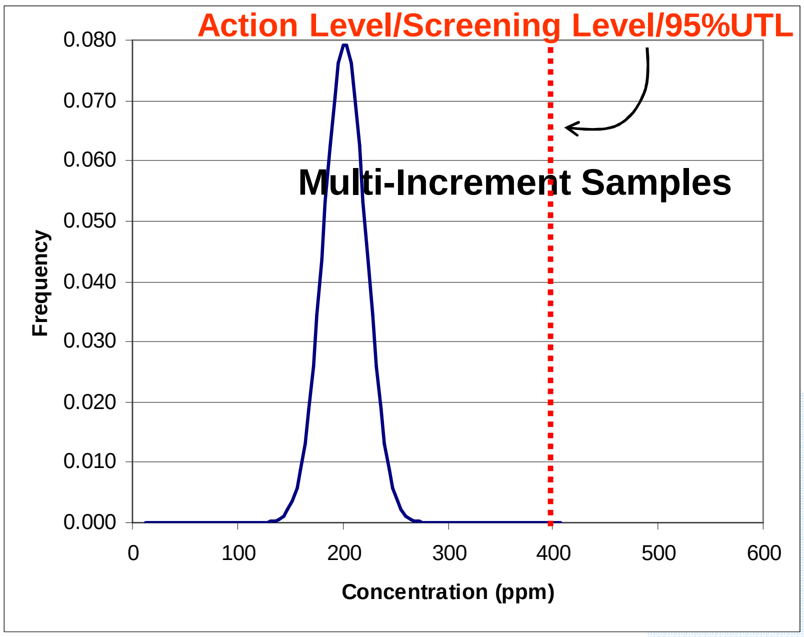
Each sample support has its own distribution of data values.

The larger the sample support, the closer to normal the distribution of data values becomes, and the fewer results appear in a distribution tail.

Explains why repeat sampling can fail to find the "hot spot"

Here's the problem:

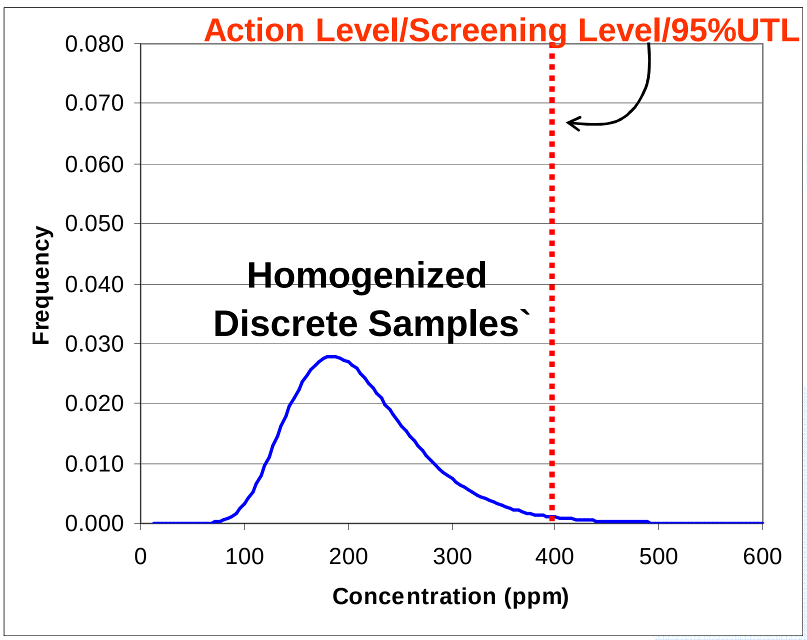
Underlying sample population distribution is tightly tied to the sample support: different sample supports have different statistical distributions



Large sample support (assume sample is properly processed for subsampling)

Here's the problem:

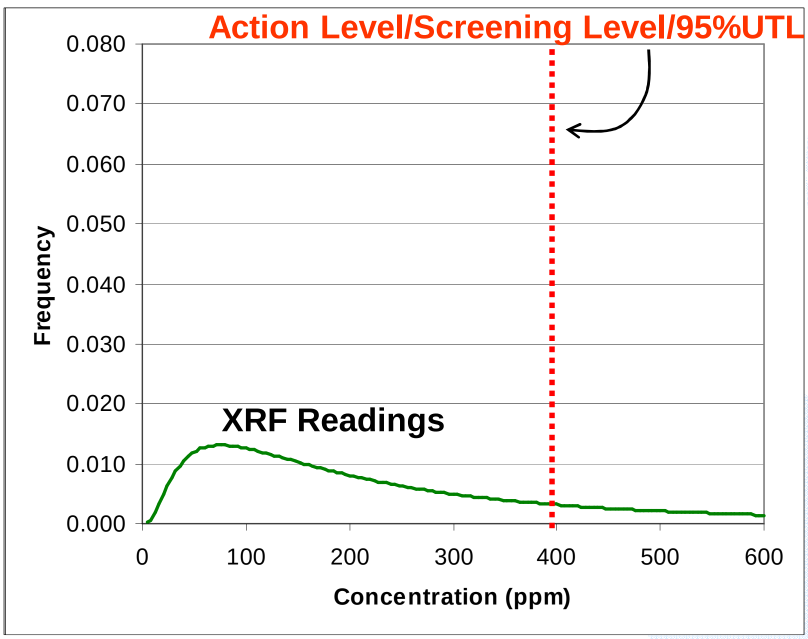
Underlying sample population distribution is tightly tied to the sample support: different sample supports have different statistical distributions



Medium sample support (assume sample is properly processed for subsampling)

Here's the problem:

Underlying sample population distribution is tightly tied to the sample support: different sample supports have different statistical distributions



Very small sample support

“Hot spot” <u>Must</u> be Defined

- Cannot scientifically defend defining a “hot spot” as <u>any</u> exceedance of an action level by a single sample result
- A scientifically defensible hot spot will be defined by concentration AND by the area/volume over which that concentration is averaged

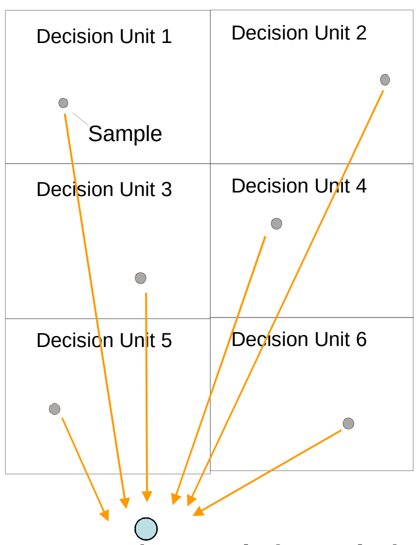
Be Careful!

- Incremental sampling may not be appropriate when evaluating the presence of contamination if the threshold is based on a discrete sample result
- When comparing a composite result to a BTV, it is essential that the background area used to develop the BTV be sampled with the same incremental-composite protocols
- Best when we are explicit about the derivation of the “decision unit” that is the basis for the decisions we are making, as well as the meaning of “hot spot”

Composite-Searching vs Incremental-Averaging

Assumption: looking for evidence of contamination across units

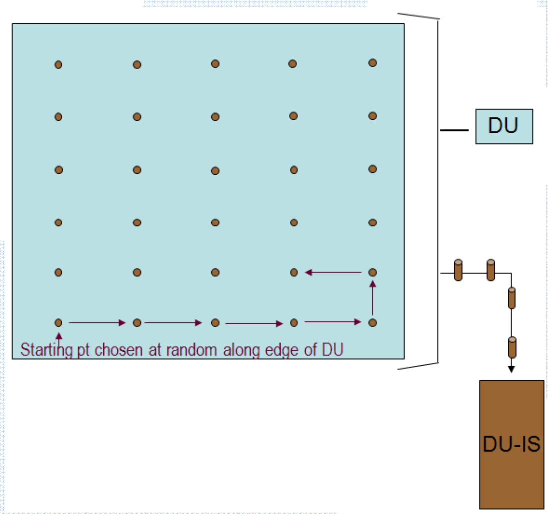
Composite-Searching



Form composite sample for analysis

Assumption: cleanup criteria averaged over decision unit

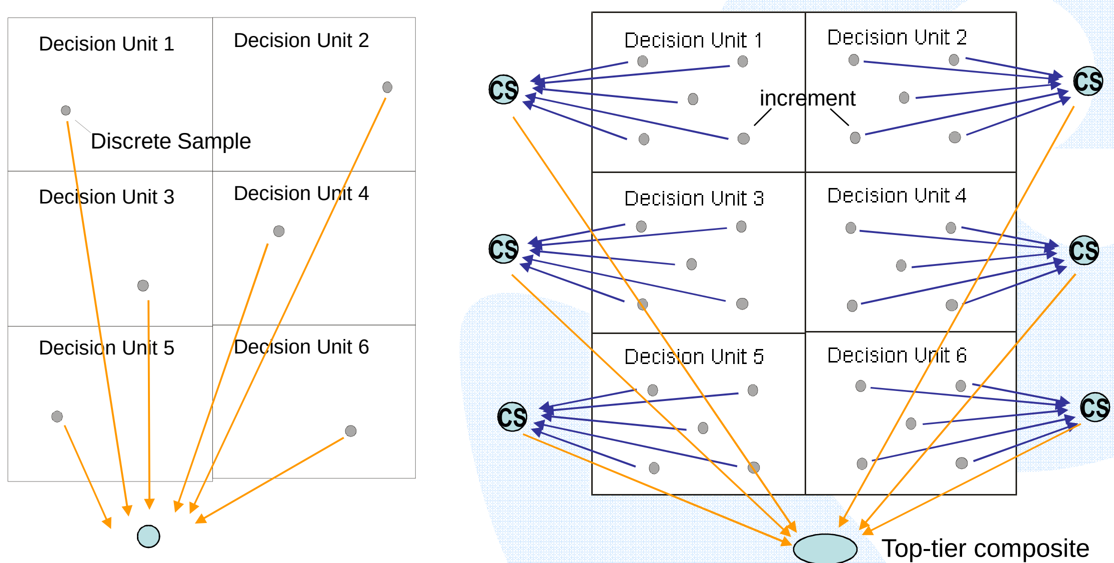
Incremental-Averaging



2 Strategies for Composite-Searching

Search-composite formed of *discrete samples*

Search-composite formed of *bottom tier composite samples (CSs)*



Form search-composite for analysis

Form search-composite for analysis

Adaptive Compositing Strategies for Searching

- Goal is to identify elevated areas
  - Looking for contamination > designated action level
- Assumptions:
  - Contamination believed to be spotty
  - Action level significantly > background levels
  - Sample acquisition/handling costs significantly < analytical costs
  - Appropriate methods exist for sample acquisition & aggregation

Adaptive Composite-Searching (cont'd)

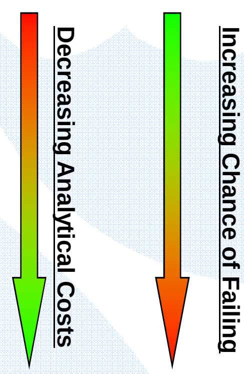
- Aggregate samples (discrete or IS) into composites for homogenization and analysis
  - Split each discrete (or IS, as the case may be)
  - Composite 1 set of splits
  - Archive the other set for re-analysis if necessary
- For the design:
  - Determine appropriate number of samples to composite (see next few slides)
  - Develop decision criteria for composites that indicate when analyses of archived splits are necessary (next)

Recipe for Adaptive Composite-Searching

- Determine appropriate number of samples to composite & resulting decision criteria; use equal masses/volumes
- Decision criteria = [(action level - background) / (# of samples in composite)] + background
- Homogenize well & split samples – use one set of splits to form composites and archive other set
- If composite result < decision criteria
  - No more sampling required
- If composite result > decision criteria
  - Analyze archived splits contributing to composite

Example Decision Criterion

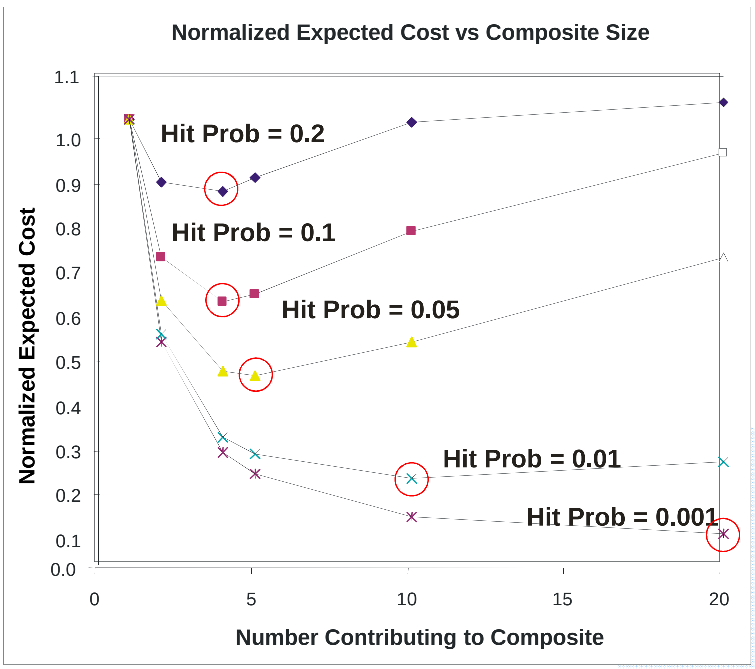
- Background: 10 ppm, Action Level: 100 ppm
- Determine decision criteria for 2-sample, 3-sample, 4-sample, 5-sample, and 6-sample composite:
  - 2-sample composite: 55 ppm
  - 3-sample composite: 40 ppm
  - 4-sample composite: 33 ppm
  - 5-sample composite: 28 ppm
  - 6-sample composite: 25 ppm
  - Don't want to get too close to background



When is Adaptive Compositing Cost-Effective?

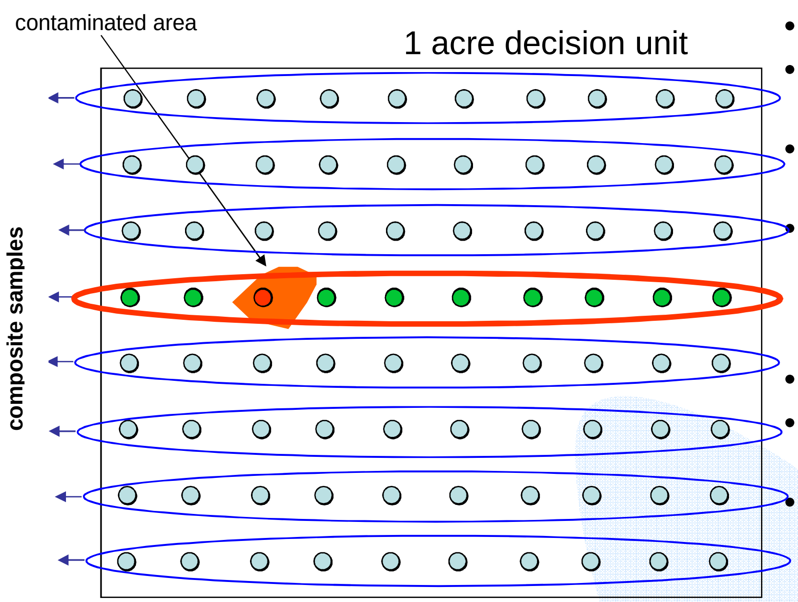
- The "spottier" contamination is, the better the performance (in contrast to discrete sampling)
- The greater the difference is between background and the action level, the better the performance
- The greater the difference between the action level and average contamination concentration, the better the performance
- Best case: no composite requires re-analysis
- Worst case: every composite requires re-analysis (will cost more than if the samples were just analyzed as discretes from the start)

Optimize Cost: How Many Samples to Composite?



- Is a function of the probability that contamination is present such that a composite will exceed the criteria
- Balance against the cost of going back & analyzing the archived increment splits when a hit occurs (which is why the lines go back up)
- The less likely it is that contamination is present, the more samples that can be composited
- The circled point identifies the ideal sample number to composite based likelihood of exceedence

Simplified Example…

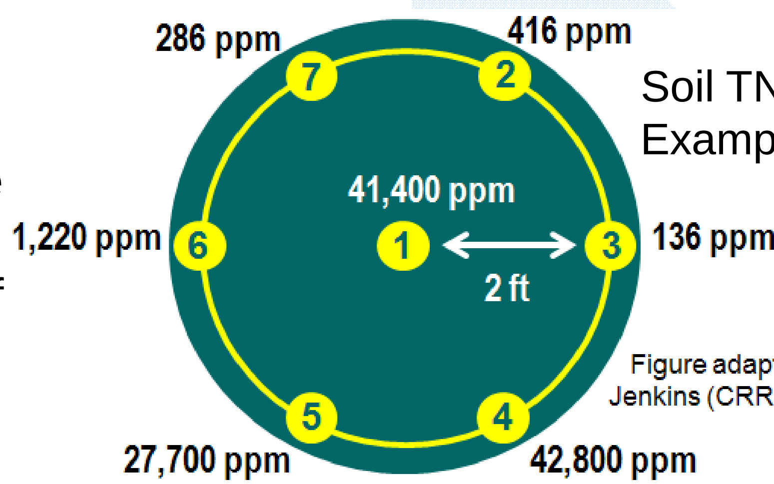


- Looking for PCBs > 50 ppm
- Could be anywhere in area of concern 1 acre in size
- Discrete sample/analyze all approach – 80 samples
- Alternative: take 80 samples and organize into 8 composites consisting of 10 samples each
- Decision criteria: **5 ppm**
- Would need to analyze 8 composites
- Plus analyze 10 archived samples from the guilty composite
- Find hotspot & calc average over all or portion of DU

More than 75% analytical cost reduction!!

A Flawed Assumption About "Hot Spots"…

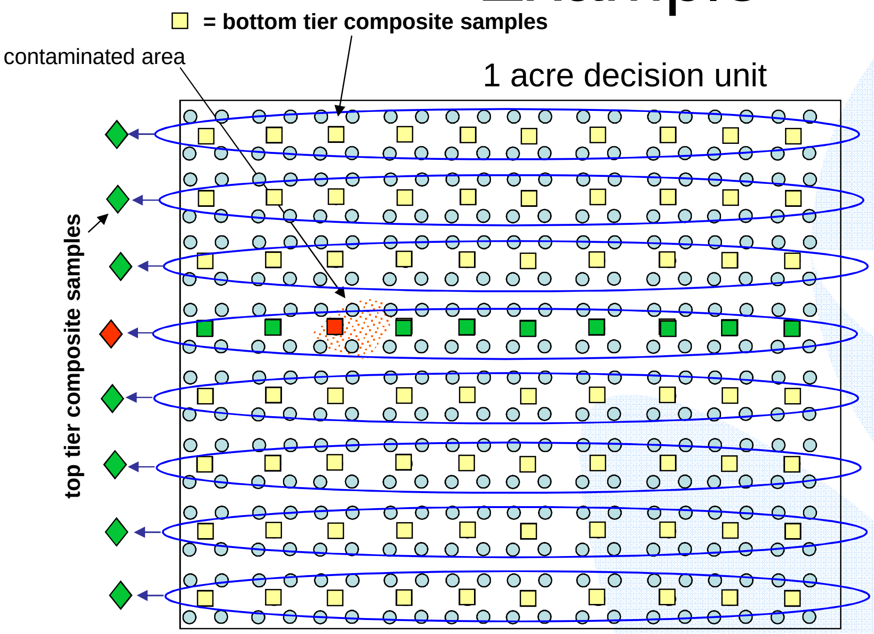
- Average concentration across seven samples is 16,000 ppm
- Suppose "hot spot" level were 5,000 ppm
- Only a 43% chance of correctly identifying the presence of this hot spot if one relied on discrete samples



Soil TNT Example

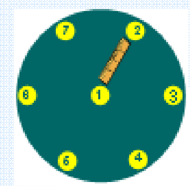
Figure adapted from Jenkins (CRREL), 1996

Two Tier Composite Search Example



- 400 soil increments ○ grouped by 5's
- Form 80 bottom-tier □ composite samples representative of 80 small areas
- 10 bottom-tier composites combined to form 8 top-tier composites, which are analyzed (◆ &◆)
- 10 archived bottom-tier splits required analysis due to the hit for a top-tier composite (◆)
- Rationale: improve likelihood of "seeing" contamination

Review Outputs from Combined Design



1. High density sampling coverage: controls short-scale heterogeneity & its problems
2. Finds hot spots: can go back & do finer delineation if desired
3. Can calculate DU average over whole DU or only selected parts
4. Provides vastly improved information for about same, or less, cost than much less informative sampling designs
5. Supports high confidence, high precision remedial designs

EPA/600/R-03/027
November 2003

Guidance for Obtaining Representative Laboratory Analytical Subsamples from Particulate Laboratory Samples

http://www.clu-in.org/download/char/epa_subsampling_guidance.pdf

Particle Size Segregation

Freshly collected soil sample in jar – no segregation by particle size

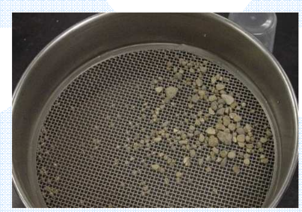
Same sample jar after "jiggling" to mimic transportation to lab - segregation evident
What if scoop off the top?

48

Photo credits: Deana Crumbling

Sample Processing Prior to Subsampling

- Obviously depends on soil type, moisture, etc.)
  - Drying: oven or air-drying
  - Disaggregation: breaking up clods
  - Sieving: separate out particle size fraction of interest
  - Grinding: mills & grinders (under some circumstances)



Gy theory guides sample handling & subsampling to produce the analytical sample

1-D Slab Cake Sample Mass Reduction

50

2-D Slab Cake Sample Mass Reduction

Field subsampling from core to sample jar

Picture from USACE-Alan Hewitt

See ITRC ISM-1, Section 2.6.6.7

Lab subsampling to reduce sample volume for further processing

Photo courtesy of Mark Bruce, TestAmerica

From EPA 600/R-03/027

From Table 8. Subsampling methods that are NOT recommended based on experimental evaluation

Method	Typical Increment Size	Sensitivity to Grouping & Segregation	Moisture Content	Correct Sampling Possible	Agreement with Calculated s_{FE}^2	Comments
Degenerate Fractional Shoveling	Medium to Large	Moderate to High	Dry to Moist	Yes, if Careful	Unlikely	Performance Tied to Lot Mass; Subject to Bias; N.R.
Rolling and Quartering	Large	High	Dry	Yes, if Careful	Usually Not Close	Highly Variable; NR
Coning and Quartering	Large	High	Dry	Yes, if Careful	Usually Not Close	Usually Biased; NR
V-Blender	N.A.	High	Dry	N.A.	Very Unlikely	Problems with GE; N.R.
Vibratory Spatula	Small	Very High	Dry	No	Not Close	Problems with GE; NR
Grab Sampler	Variable	Very High	Dry to Moist	No	Not Close	Biased and Variable; N.R.

When is Compositing Potentially Not Appropriate?

- Contaminants subject to loss during sample handling (e.g., mercury, VOCs)
- Analytical costs are minimal (e.g., XRF)
- Sample acquisition costs are significant (e.g., subsurface sampling)
- No provisions are made for proper laboratory sample handling/preparation

Any Questions?