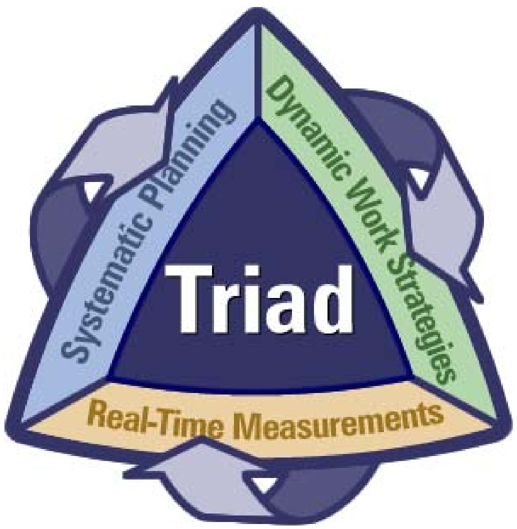
Sampling Design
Avoiding Pitfalls in Environmental Sampling

Part 2



U.S. EPA Technical Session Agenda

- Welcome
- Understanding: Where does decision uncertainty come from?
- Criteria: You can't find the answer if you don't know the question!
- Pitfalls: How to lie (or at least be completely wrong) with statistics…

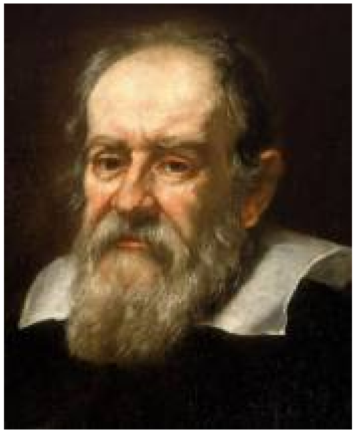
- Solution Options: Truth serum for environmental decision-making

Solution Options:

Truth serum for environmental decision-making

"All truths are easy to understand once they are discovered; the point is to discover them."

Galileo



Triad Data Collection Design and Analysis Built On:

- **Planning systematically** (Conceptual Site Models [CSM])
- **Improving representativeness**
- **Increasing information** available for decision-making via real-time methods
- **Addressing the unknown** with dynamic work strategies

Systematic Planning Addresses:

Planning Systematically

- Obscure the question…
- The representative sample that isn't…
- Pretend the world is normal…
- Assume we know when we don't…
- Ignore short-scale heterogeneity…
- Miss the forest because of the trees…
- Regress instead of progress…
- Statistical dilution is the solution…
- Worship the laboratory…

Systematic Planning and Data Collection Design

Planning Systematically

- Systematic planning defines decisions, decision units, and sample support requirements
- Systematic planning identifies sources of decision uncertainty and strategies for uncertainty management
- Clearly defined cleanup standards are critical to the systematic planning process
- CSMs play a foundational role

CSMs Articulate Decision Uncertainty

Planning Systematically

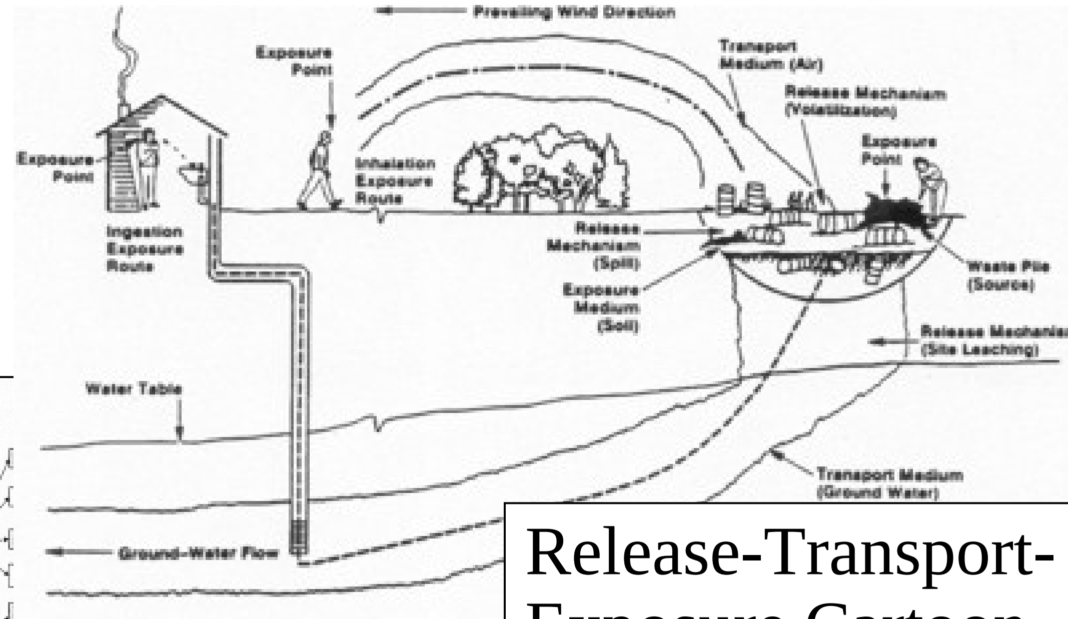
- CSM captures understanding about site conditions.
- CSM identifies uncertainty that prevents confident decision-making.
- A well-articulated CSM serves as the point of consensus about uncertainty sources.
- Data collection needs and design flow from the CSM:
  - » Data collection to reduce CSM uncertainties
  - » Data collection to test CSM assumptions
- The CSM is living…as new data become available, it is incorporated and the CSM matures.

A CSM Organizes Information

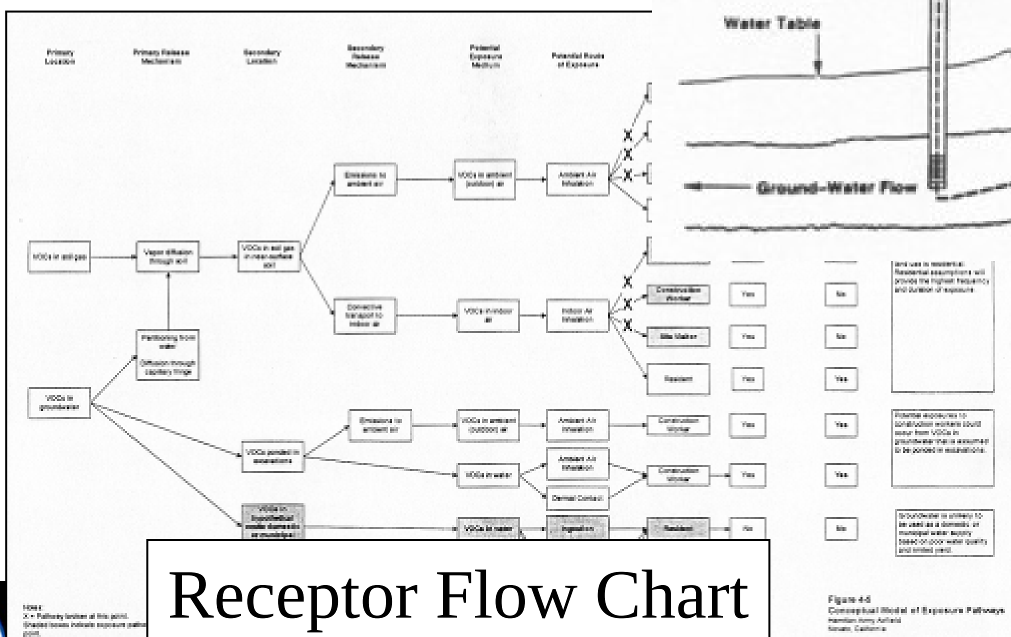
Planning Systematically

Text
(simple scenarios)

Computer Model



Release-Transport-Exposure Cartoon

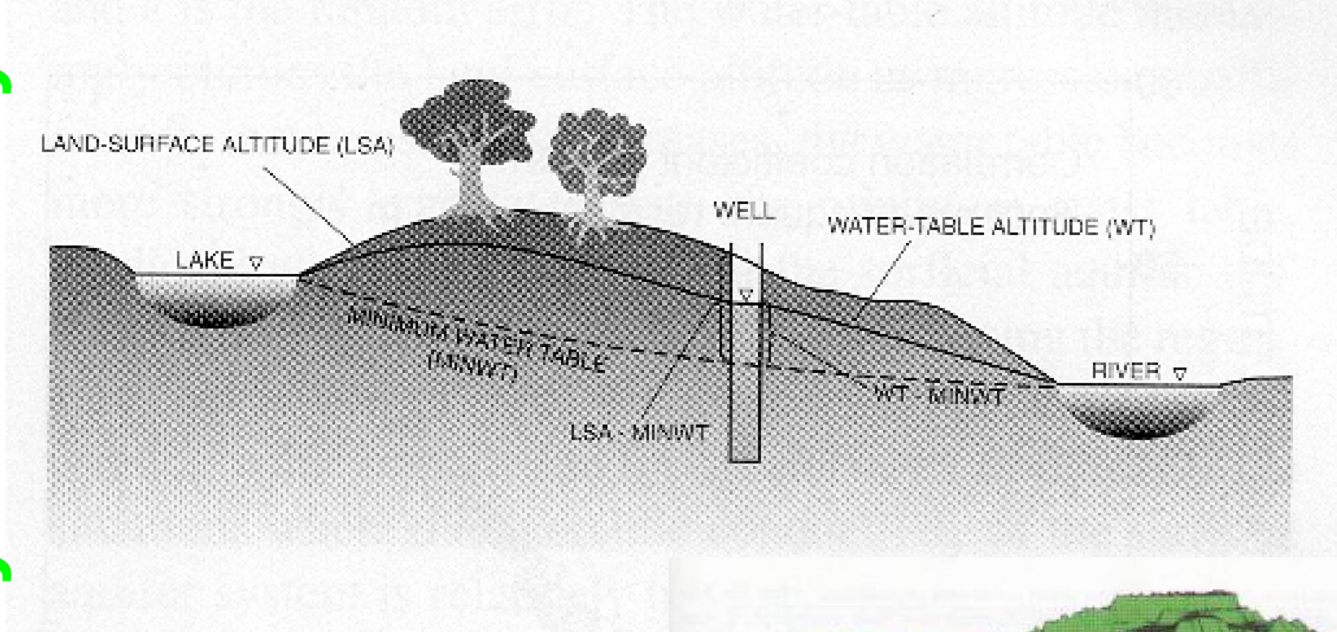


Receptor Flow Chart

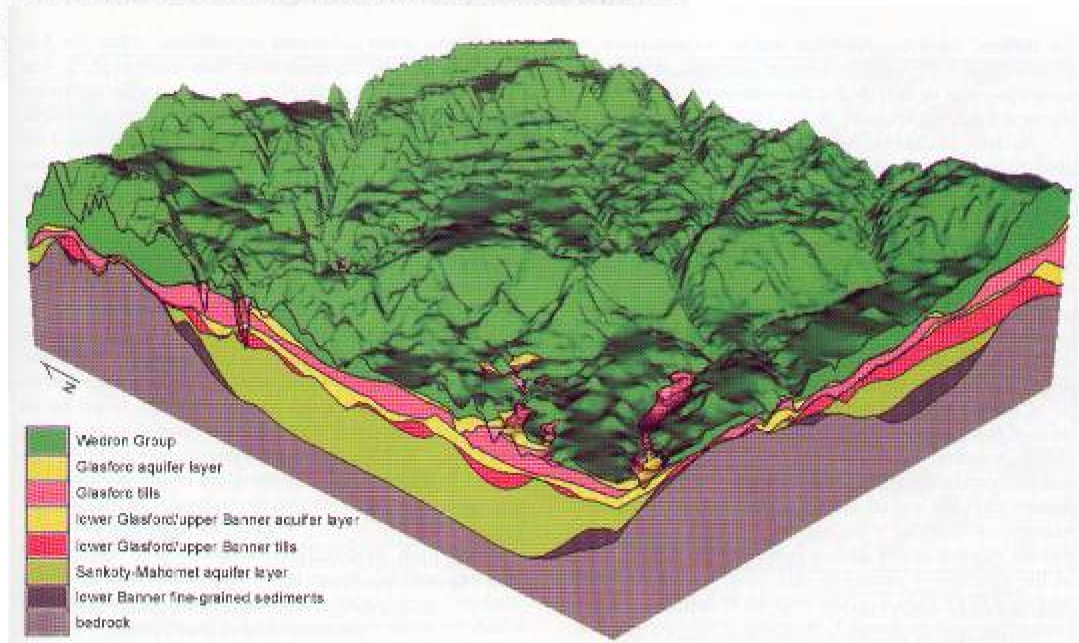
2-D Cross Section

Other CSM Representations

Planning Systematically



2-D Hydrology Model



3-D Geology Model

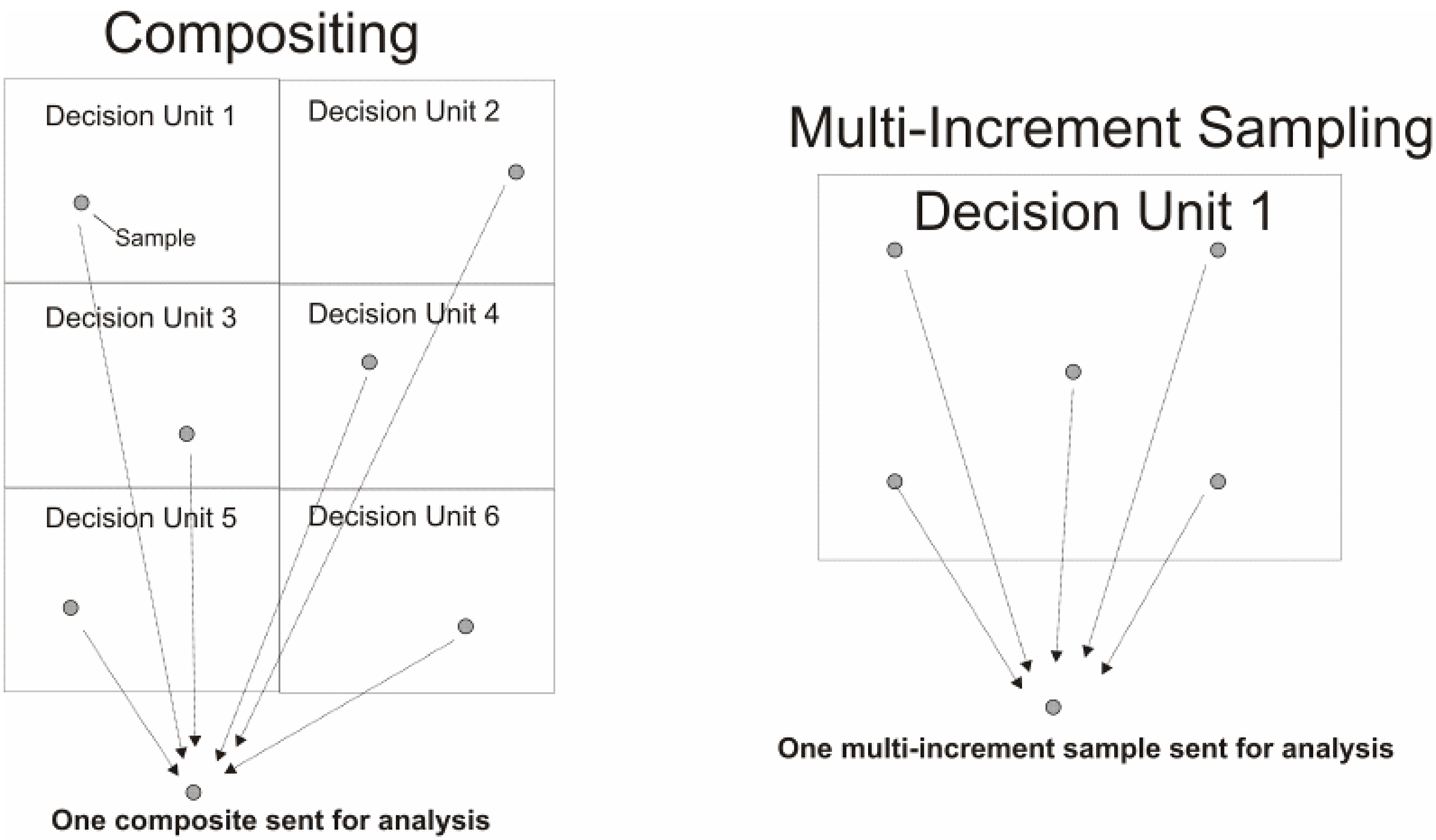
Improving Data Representativeness Addresses:

- Obscure the question…
- The representative sample that isn't…
- Pretend the world is normal…
- Assume we know when we don't…
- Ignore short-scale heterogeneity…
- Miss the forest because of the trees…
- Regress instead of progress…
- Statistical dilution is the solution…
- Worship the laboratory…

Improving Data Representativeness

- Sample support
  - » Matching sample support with decision needs
  - » Field of view for in situ analyses
- Controlling within-sample heterogeneity
  - » Appropriate sample preparation important (see EPA EPA/600/R-03/027 for additional detail)
  - » Uncertainty effects quantified by appropriate sub-sample replicate analyses
- ***Controlling short-scale heterogeneity***
  - » Multi-increment sampling
  - » Aggregating in situ measurements

Multi-Increment Sampling vs. Compositing



Asumption: cleanup criteria averaged over decision unit

Multi-Increment Sampling vs. Compositing

- Multi-increment sampling: a strategy to control the effects of heterogeneity cost-effectively
- Compositing: a strategy to reduce overall analytical costs when conditions are favorable

Improving Representativeness

Multi-Increment Sampling

- Used to cost-effectively suppress short-scale heterogeneity
- Multiple sub-samples contribute to sample that is analyzed
- Sub-samples systematically distributed over an area equivalent to or less than decision requirements
- Effective when the cost of analysis is significantly greater than the cost of sample acquisition

How Many Increments?

- Enough to bring short-scale heterogeneity under control relative to other sources of error
- Practical upper limit imposed by homogenization capacity
- Level of heterogeneity present best estimated with a real-time technique
  - » At least 10 measurements to get an estimate of variability at contaminant level of interest
  - » Variability estimate can be used to determine how many increments should be used

Increasing Information via Real Time Methods

Increasing Information

- Obscure the question…
- The representative sample that isn't…
- Pretend the world is normal…
- Assume we know when we don't…
- Ignore short-scale heterogeneity…
- Miss the forest because of the trees…
- Regress instead of progress…
- Statistical dilution is the solution…
- Worship the laboratory…

What Are Real Time Methods?

Increasing Information

- Any analytical method that provides data quickly enough to affect field work while underway
- Typically are field-based, but not necessarily so
- Field-based methods are cheaper and faster
  - » Allow increased data densities
  - » Allow real-time decision-making and optimization of the sampling approach
  - » Typically higher detection limits, or greater analytical error, or less specificity, or more limited range of analytes
- Typically produce “collaborative data sets”
  - » Collaborative data sets are powerful!!
  - » Multiple lines of evidence = “weight of evidence”
  - » Control multiple error sources
  - » Result: increased confidence in the CSM, better decisions

Examples of Common Field-Based Methods

Increasing Information

Technology	Matrix	Data Provided
X-ray fluorescence (XRF)	Soil	Metals
Immunoassay test kits	Water, Soil	SVOCs (PAH, pest., PCB)
UV methods (UVF, UV lamp)	Water, Soil	TPH, PAH, DNAPL
Misc. colorimetric kits	Water, Air	Water Quality, Toxic Gas
Direct push sensors (MIP, DSITMS, LIF)	Water, Soil	VOCs, TPH, DNAPL
Geophysical tools	Soil	Sources, Pathways
Field GC and GC/MS	Water, Soil	VOCs

Collaborative Data Sets Address Analytical and Sampling Uncertainties

Increasing Information

Cheaper/rapid (lab? field? std? non-std?) **analytical methods**

↓

Targeted high density sampling

↓

Manages CSM & sampling uncertainty

Costlier/rigorous (lab? field? std? non-std?) **analytical methods**

↓

Low DL + analyte specificity

↓

Manages analytical uncertainty

Collaborative Data Sets

Regression Analysis is Not Always Appropriate for Collaborative Data Sets

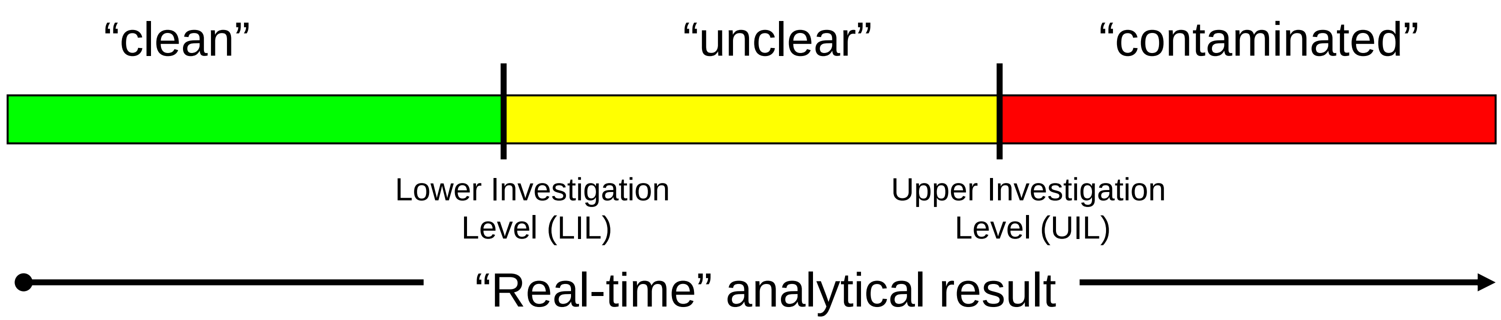
Increasing Information

- Beryllium/LIBS (laser-induced breakdown spectroscopy) example illustrates:
  - » Site with surficial beryllium contamination
  - » Experimental backpack LIBS system used to generate a lot of characterization data
  - » Much smaller set of laboratory analysis for a subset of LIBS locations also available
  - » Question: Where's the contamination?

Non-Parametric Analysis Can Be a Useful Alternative to Regressions

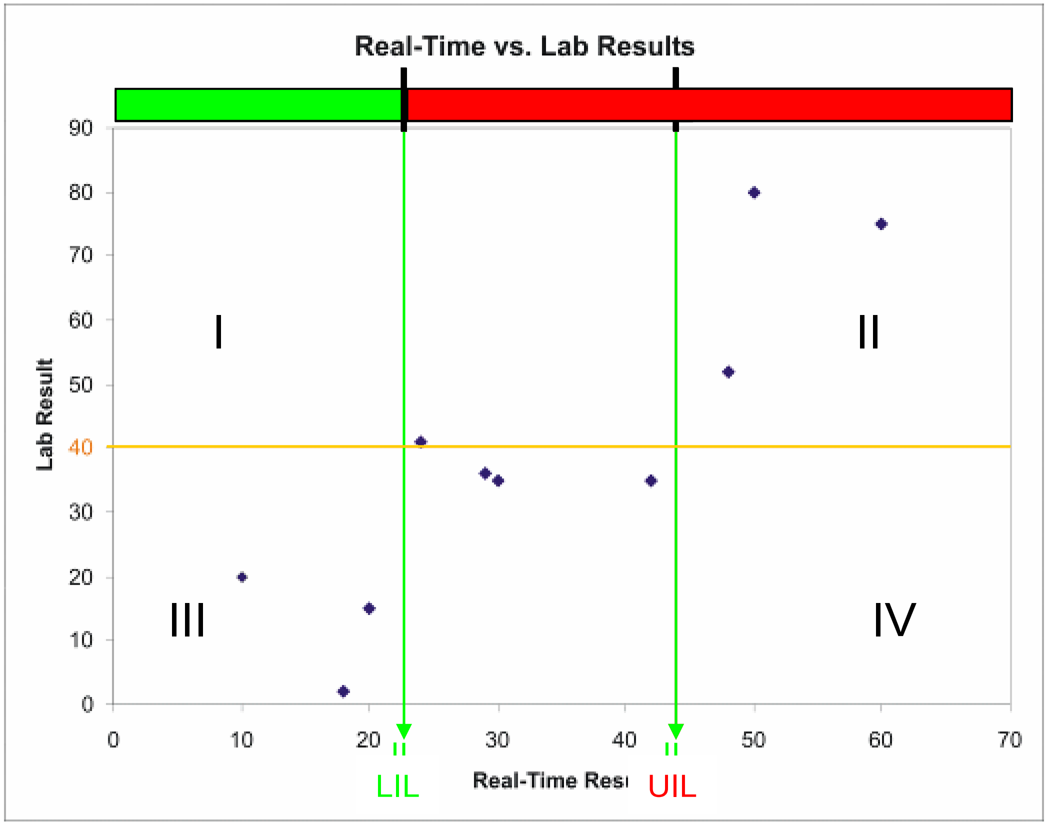
Increasing Information

- Decision focus is often yes/no
  - » Is contamination present at levels of concern?
  - » Should a sample be sent off-site for more definitive analysis?
- Goal is to identify investigation levels for a real-time method that will guide decision-making
  - » Lower investigation level (LIL) for real-time result below which we are confident contamination is not present
  - » Upper investigation level (UIL) above which we are confident contamination is present



Hypothetical Example

Increasing Information



- I: False Clean
- II: Correctly Identified Contaminated
- III: Correctly Identified Clean
- IV: False Contaminated

- I/(I+II)*100: % of contaminated samples missed by LIL (false clean rate)
- I/(I+III)*100: % of "clean" samples that are contaminated
- IV/(II+IV)*100: % of "contaminated" samples that are clean
- IV/(III+IV)*100: % of clean samples above the LIL (false contaminated rate)

False Clean Rate: 0% **False Contaminated Rate: 0%**

Addressing the Unknown through Dynamic Work Strategies

Addressing the Unknown

- Obscure the question…
- The representative sample that isn't…
- Pretend the world is normal…
- Assume we know when we don't…
- Ignore short-scale heterogeneity…
- Miss the forest because of the trees…
- Regress instead of progress…
- Statistical dilution is the solution…
- Worship the laboratory…

Addressing the Unknown through Dynamic Work Strategies (continued)

Addressing the Unknown

- Optimizing data collection design
  - » Strategies for testing CSM assumptions and obtaining data collection design parameters on-the-fly
- Adaptive analytics
  - » Strategies for producing effective collaborative data sets
- Adaptive compositing
  - » Efficient strategies for searching for contamination
- Adaptive sampling
  - » Strategies for estimating mean concentrations
  - » Strategies for delineating contamination

Addressing the Unknown

Optimizing Data Collection Design

- How many increments should contribute to multi-increment samples?
- What levels of contamination should one expect?
- How much contaminant concentration variability is present across decision units?
- What kinds of performance can be expected from field methods?

Much of this falls under the category of "Demonstration of Methods Applicability"

Example: XRF Results Drive Measurement Numbers

Addressing the Unknown

- Applicable to in situ and bagged sample XRF readings
- XRF results quickly give a sense for what levels of contamination are present
- Measurement numbers can be adjusted accordingly:
  - » At background levels or very high levels, fewer
  - » Number maximum when results are in range of action level
  - » Particularly effective when looking for the presence or absence of contamination above/below an action level within a sample or within a decision unit

XRF Example (cont.)

Addressing the Unknown

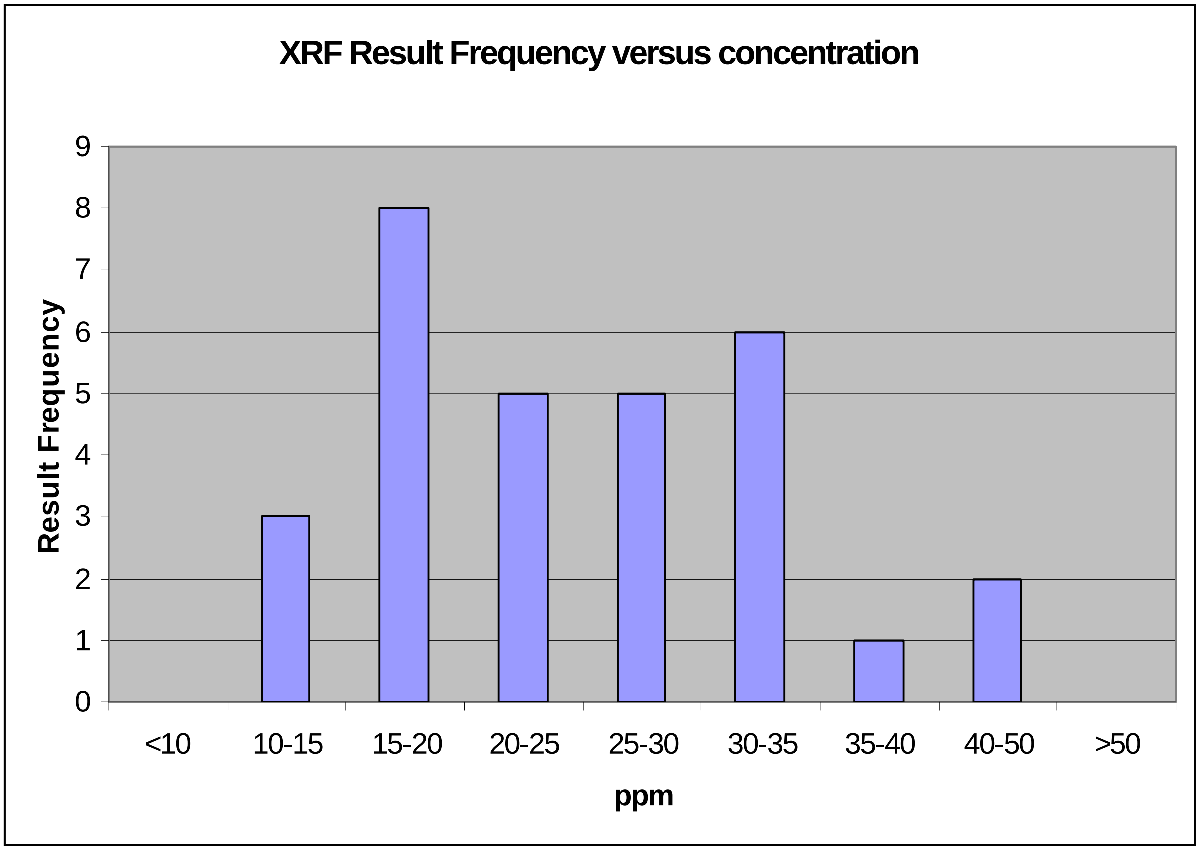
- Bagged samples, measurements through bag
- Need decision rule for measurement numbers for each bag
- Action level: 25 ppm
- 3 bagged samples measured systematically across bag 10 times each
- Average concentrations: 19, 22, and 32 ppm
  - » 30 measurements total

Addressing the Unknown

Example (continued)

Simple Decision Rule:

- If 1st measurement less than 10 ppm, stop, no action level problems
- If 1st measurement greater than 50 ppm, stop, action level problems
- If 1st measurement between 10 and 50 ppm, take another 3 measurements from bagged sample



Adaptive Analytics

Addressing the Unknown

- Goal is to identify elevated areas or delineate contamination
- Assumptions:
  - » Two methods, one cheap/less accurate, one expensive/“definitive”
  - » Investigation levels (i.e., LIL and UIL) can be derived for cheaper, real-time data
- High density real-time data used to screen out areas that are obviously contaminated, or obviously clean
- Fixed laboratory analyses target locations where real-time results were ambiguous
- Design requires determining appropriate real-time investigation levels (e.g., LIL and UIL)

Adaptive Compositing Strategies

- Goal: Identify elevated areas (or clear areas of contamination concerns)
- Assumptions:
  - » Contamination is believed to be spotty
  - » Action level is significantly greater than background levels
  - » Sample acquisition/handling costs are significantly less than analytical costs
  - » Appropriate methods exist for sample acquisition and aggregation
- Aggregate samples (single or multi-increment) into larger composites
- Design requires determining the appropriate number of samples to composite, and developing decision criteria for composites that indicate when analyses of contributing multi-increment samples are necessary

Recipe for Adaptive Compositing:

- Determine appropriate number of samples to composite and resulting decision criteria.
- Decision criteria = (action level - background)/(# of samples in composite) + background.
- Sample and split samples. Use one set of splits to composite and save other set.
- If:
  - » composite result < decision criteria, done.
  - » composite result > decision criteria, analyze splits contributing to composite.

Example Decision Criteria

Addressing the Unknown

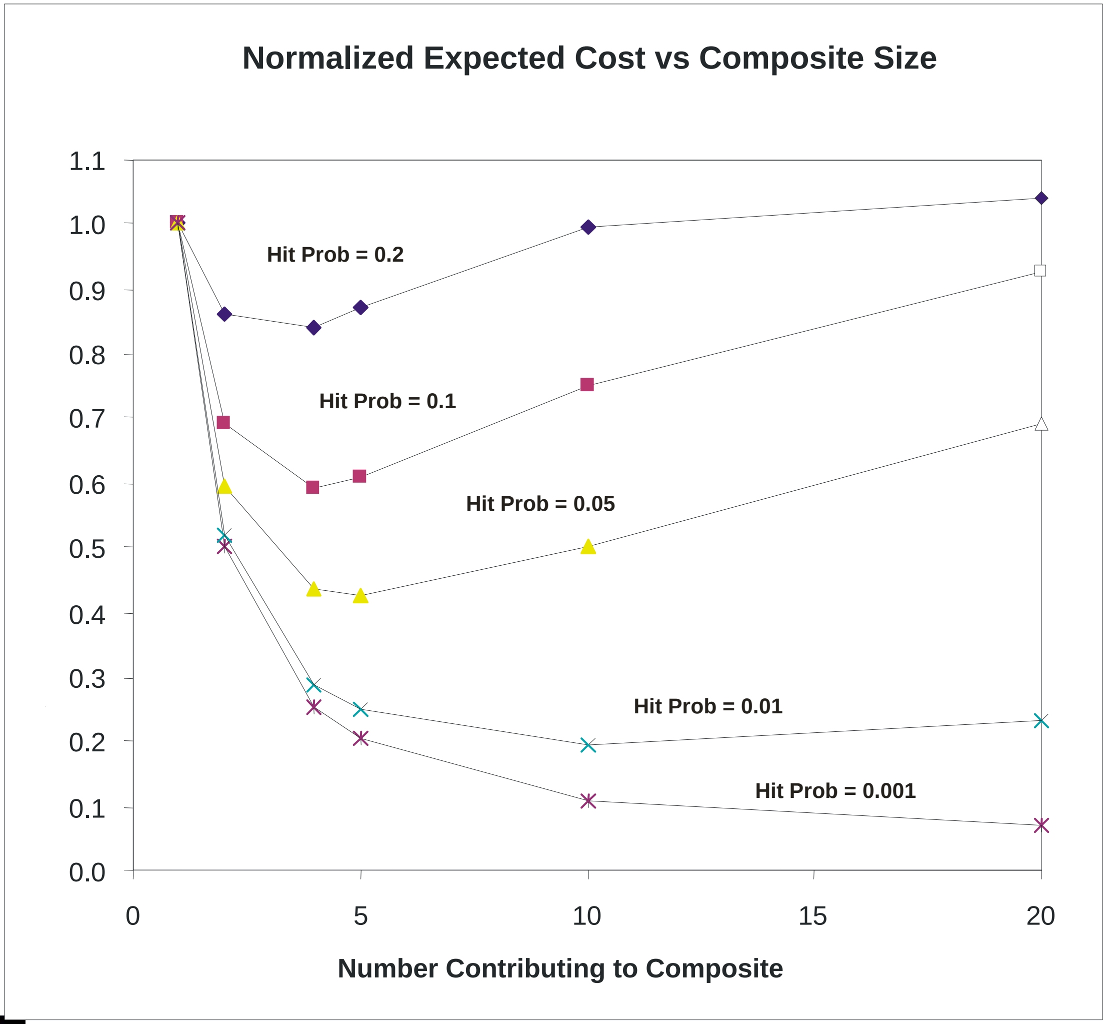
- Background: 10 ppm; Action Level: 100 ppm
- Determine decision criteria for 2-sample, 3-sample, 4-sample, 5-sample, and 6-sample composite:
  - » 2-sample composite: 55 ppm
  - » 3-sample composite: 40 ppm
  - » 4-sample composite: 33 ppm
  - » 5-sample composite: 28 ppm
  - » 6-sample composite: 25 ppm

When is Adaptive Compositing Cost-Effective?

Addressing the Unknown

- The "spottier" contamination is, the better the performance
- 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

How Many Samples to Composite?



- A function of the probability of contamination being present
- The less likely contamination is present, the larger the number of samples to composite
- Graph at left shows the case when one has 20 sampled locations

Barnard's Sequential t Test

- This statistical test is the adaptive or sequential alternative to a one-sample Student t test, used to determine if the mean concentration in a decision unit is above or below a cleanup standard.
- Sequential sampling-data evaluation-resampling… continue until the calculated average concentration is statistically either above the cleanup standard (site is "dirty"), or below the lower bound of the statistical gray region (site can be called "clean")
- Implementation requires some means for computing the required statistics. VSP has a Sequential t Test module that can do so.
- For reference on "gray region:" see EPA QA/G-4 2006, p. 53

Barnard's Sequential t Test (continued)

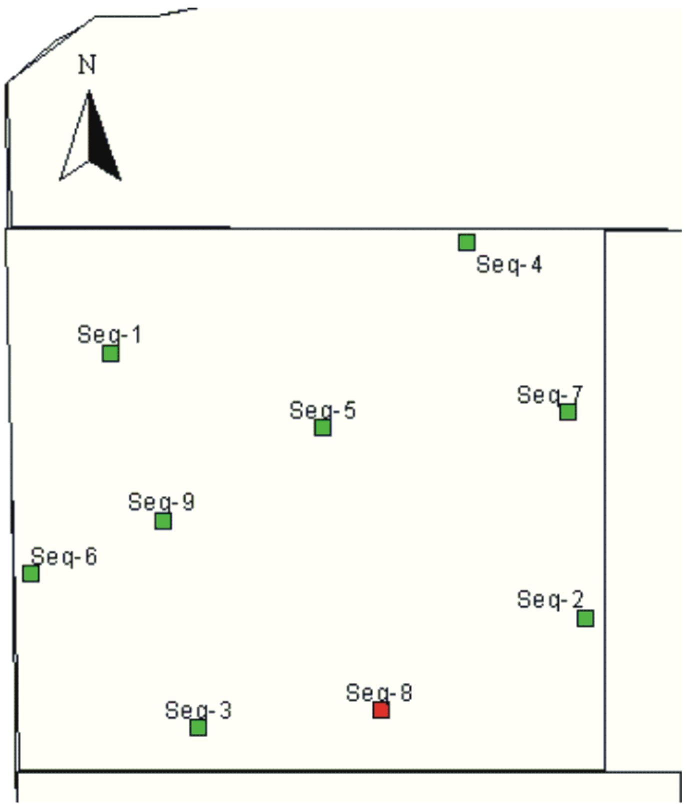
Addressing the Unknown

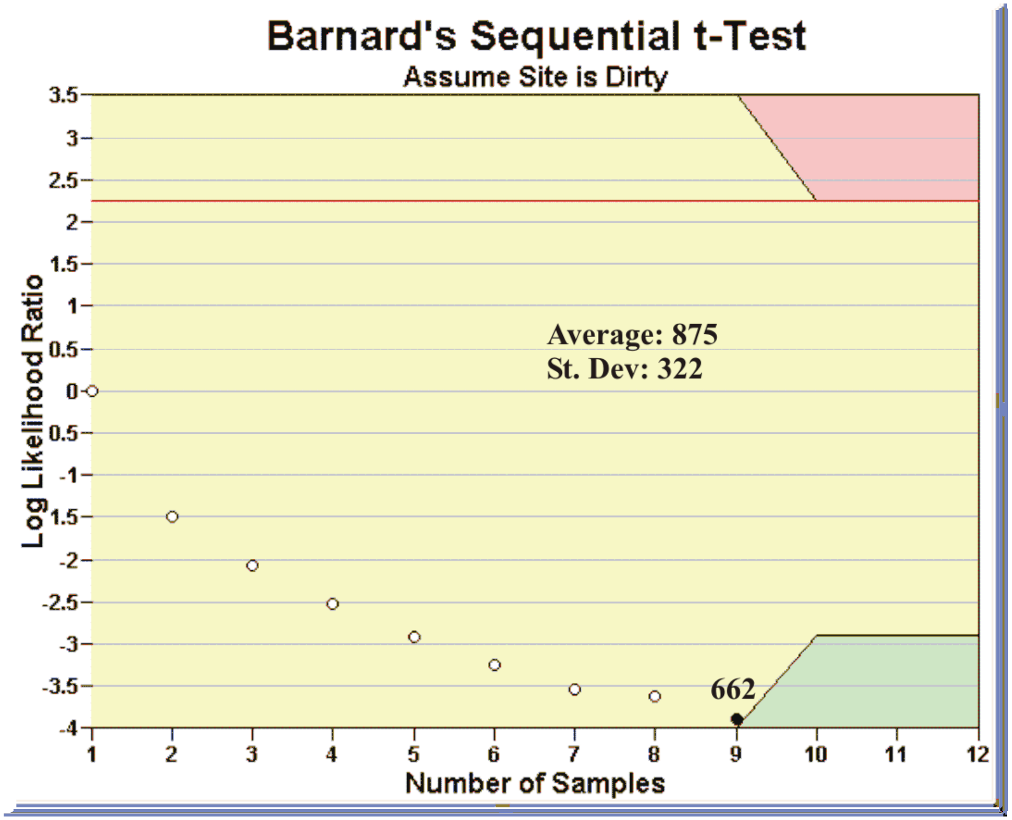
- GOAL: Determine if mean concentration is above or below the conc. standard for a decision unit.
- ASSUMPTIONS:
  - » Underlying contamination is normally distributed.
  - » You can specify some cut-off value (below the AL) that represents "definitely clean." In other words, if the average conc falls below the cut-off value, sampling can stop & the decision unit declared "clean" (this sets the lower bound of the gray region)
  - » You can specify acceptable decision error (false clean & false dirty) rates.
  - » A real-time quantitative method exists to provide unbiased data.
  - » At least 10 samples will be collected if the degree of site heterogeneity (measured by variance or SD) is unknown.
  - » Have some means for selecting the next sample locations in the sequence (e.g., adaptive fill techniques)

Sequential t Test Progress

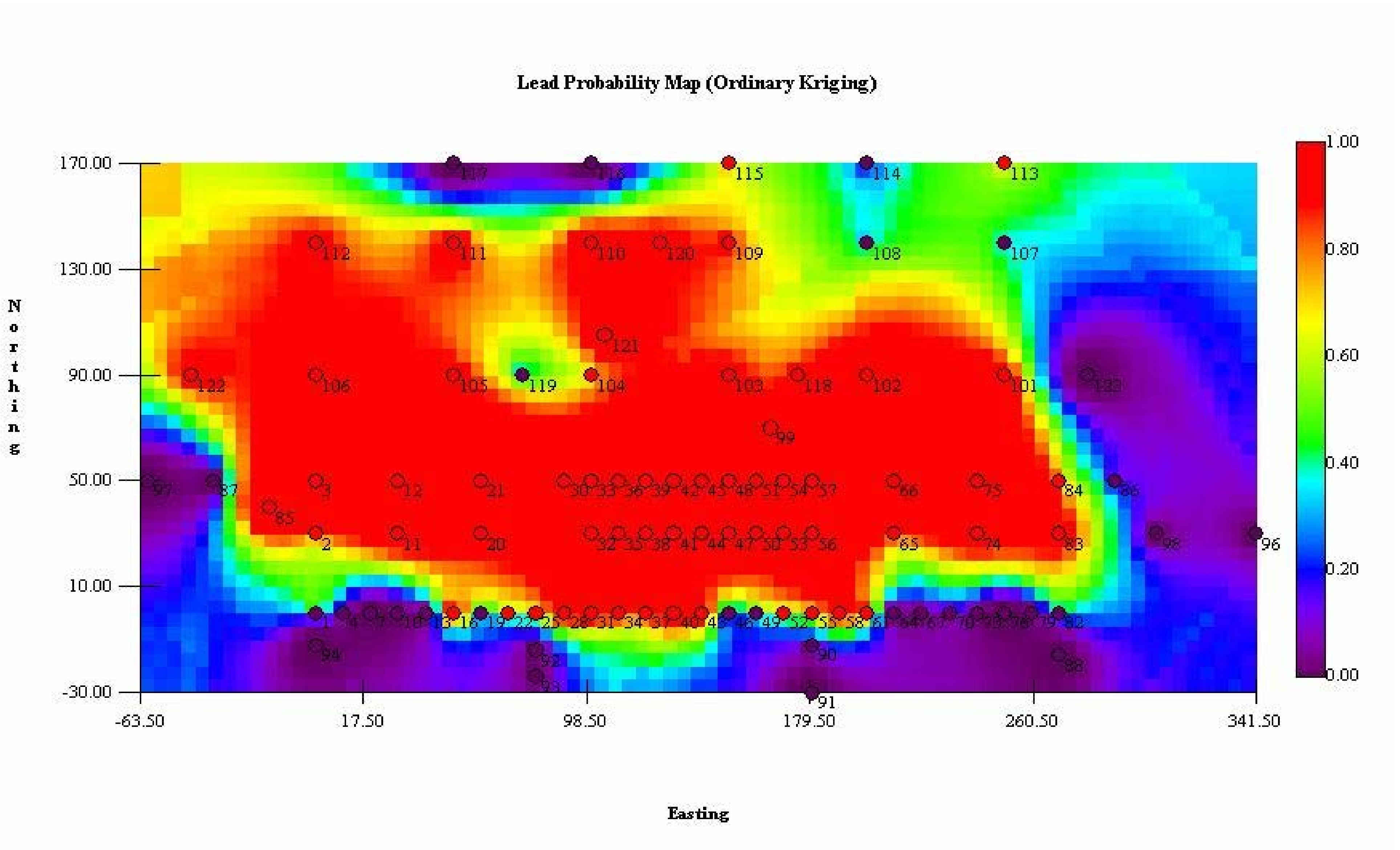
Addressing the Unknown

Cleanup Requirement: 1,500 ppm averaged over decision unit





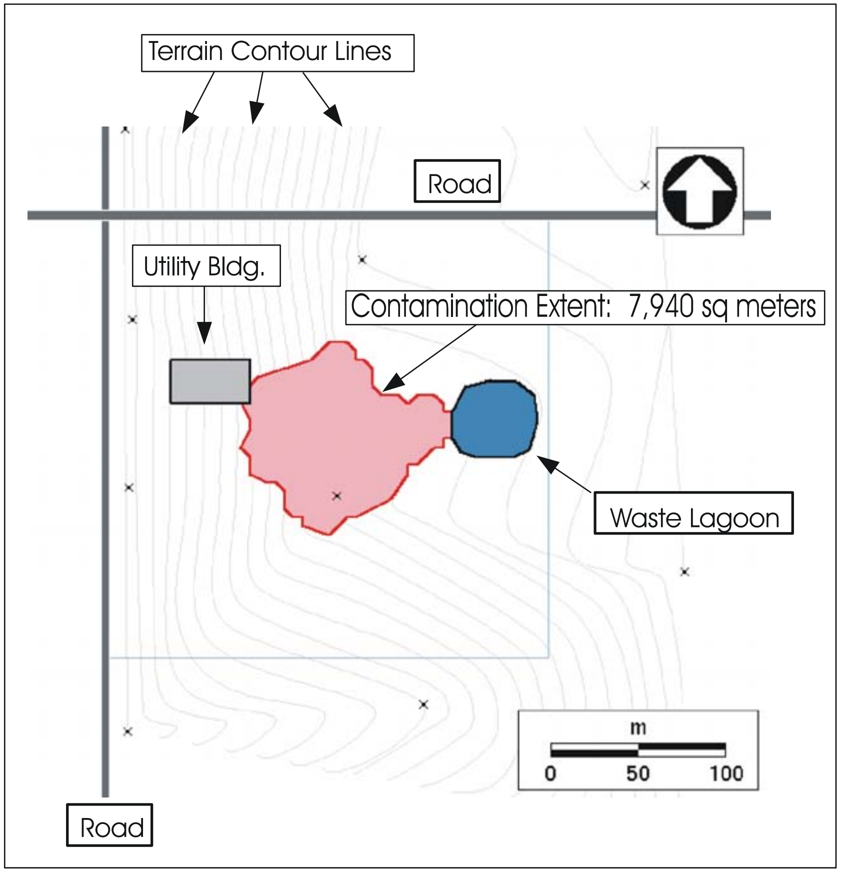
Probabilistic Decision Making using a Geostatistical Model for Soil Volumes > 250 ppm Pb



Adaptive Sampling: Delineating Surface Contamination Event

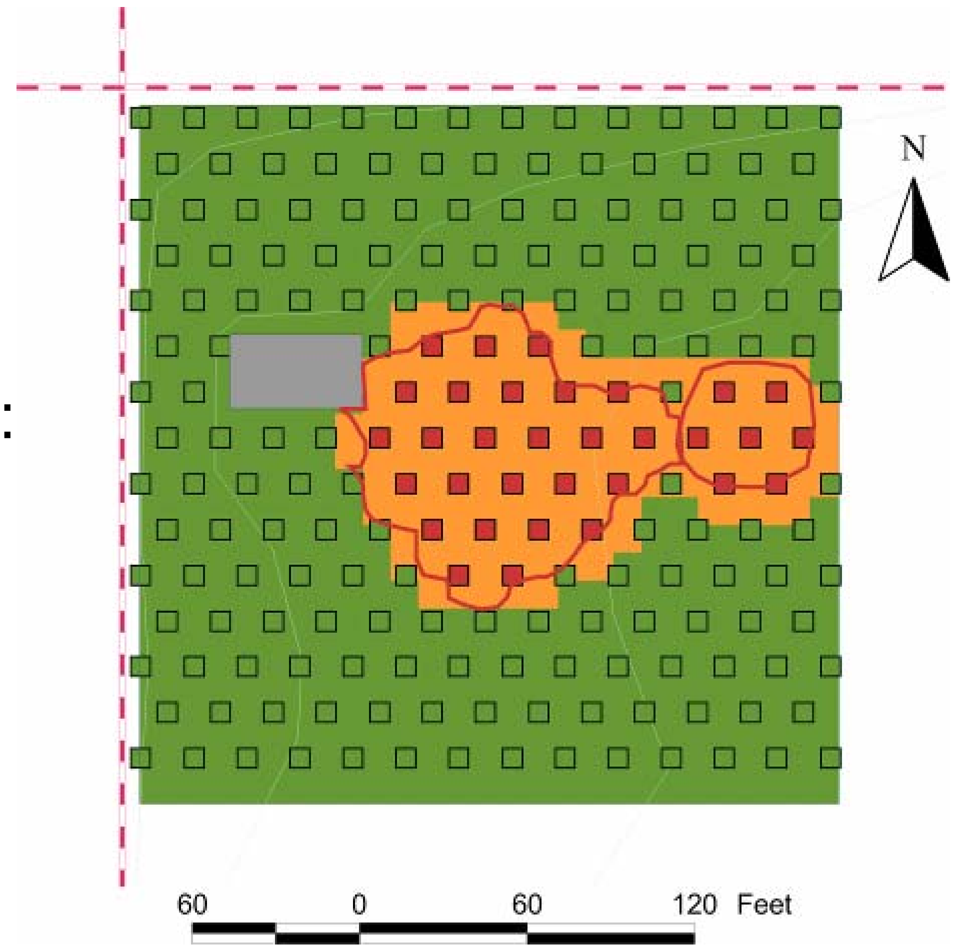
Addressing the Unknown

- Surface soil contamination problem resulting from lagoon spillage
- 10,000 sq m actually contaminated, an area unknown to the responsible party
- Soft information available for the site includes:
  - Slope of land
  - Location of barriers to flow
  - Location of source
- Owner will remediate anything with greater than 10% chance of being contaminated



Traditional Approach Results

- Triangular gridded program
- 203 samples allocated
- Observed error rates:
  - » Missed contamination: 0 ft^2
  - » Incorrectly excavated clean: 3,500 ft^2 (35% over-excavation)

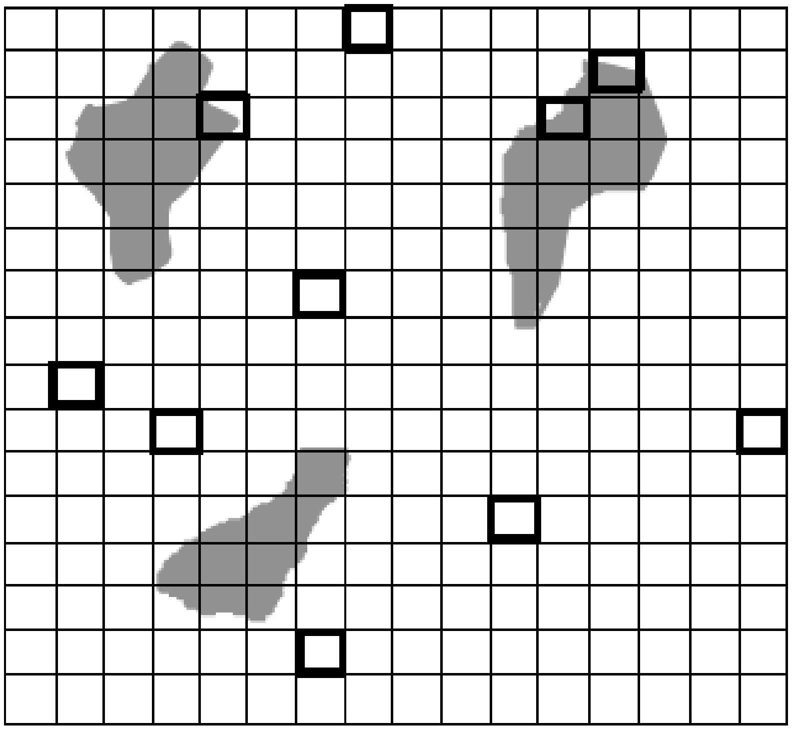


Adaptive Cluster Sampling

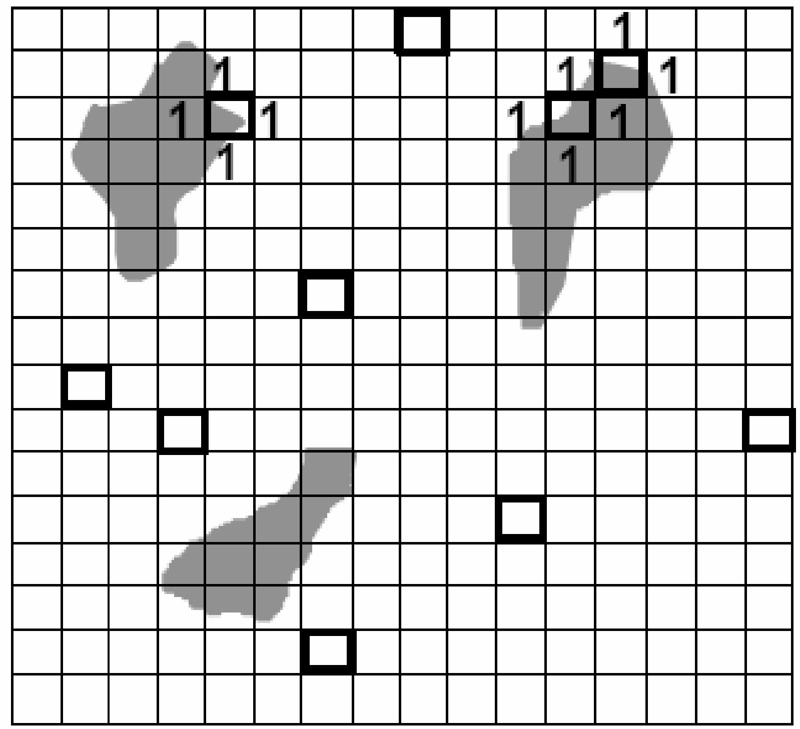
- GOAL: Determine average contaminant concentration over an area & delineate contamination footprints if any are found.
- Assumptions:
  - » The underlying distribution is normally distributed
  - » Contamination likely has a well-defined footprint
  - » Have quantitative, unbiased real-time analytics
  - » Can designate what concentration constitutes a hot spot requiring delineation
  - » Can lay a master grid over the area that encompasses all potential sampling points

Adaptive Cluster Sampling (continued)

Initial grid sampling (10 random samples)



1st batch of adjacent units

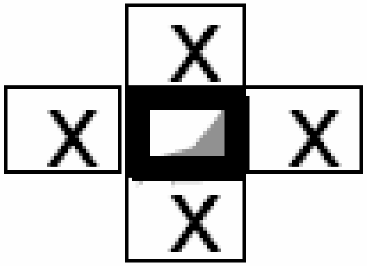


From EPA QA/G5S (Dec 2002) page 107, Figure 9-1

Adaptive Cluster Sampling (continued)

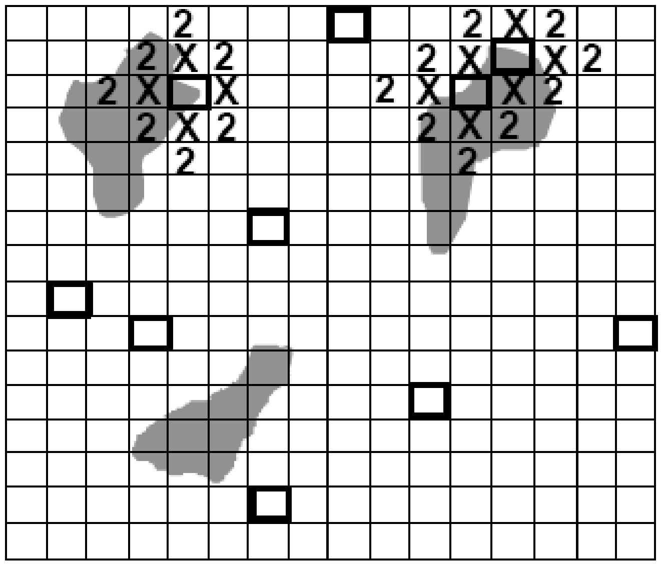
Addressing the Unknown

- Requires initial grid. The number of grid nodes to be sampled in the 1st round are determined based on the number needed to estimate a simple mean.
- Any contamination found is surrounded by samples from adjacent nodes. Sampling continues until any contamination encountered is surrounded by samples with results below the designated contaminant level.

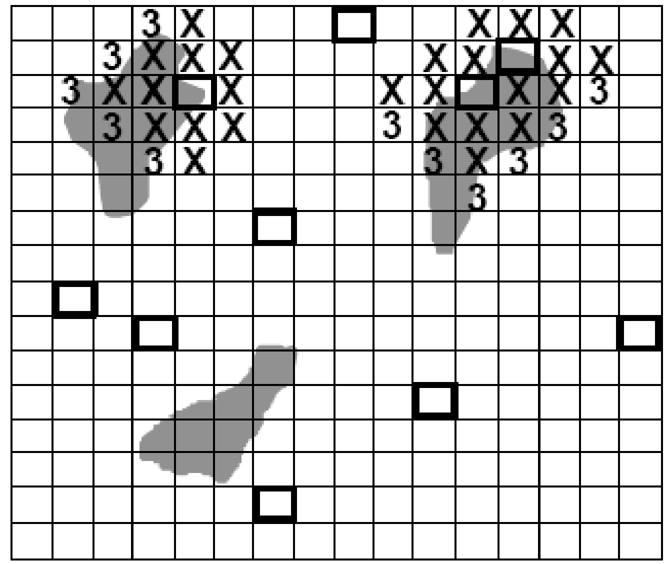


- When sampling is done, estimating the mean concentration requires a more complicated computation because of the biased sampling.
- Available in VSP.

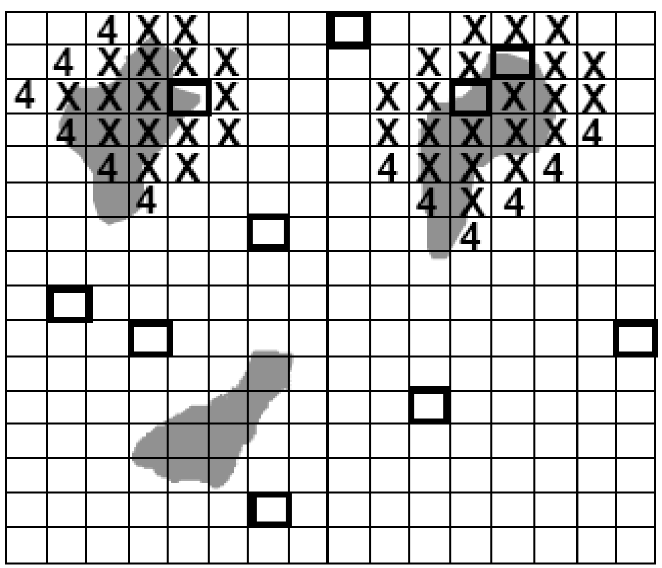
Adaptive Cluster Sampling (continued)



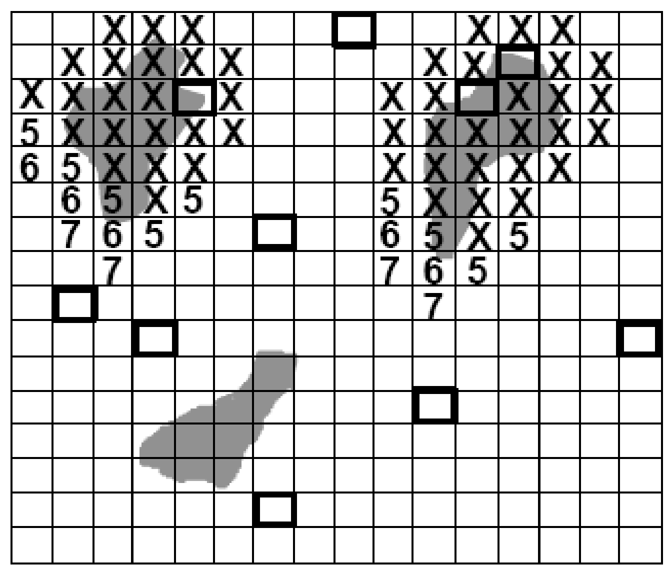
(c) Second batch of adjacent units



(d) Third batch of adjacent units



(e) Fourth batch of adjacent units



(f) Final batch of adjacent units

From EPA QA/G5S (Dec 2002) page 107 Figure 9-1

Recipe for Adaptive Cluster Sampling

- Lay master grid over site.
- Start with an initial set of gridded samples, either determined by hot spot detection design or by design to estimate conc mean.
- For every sample that is a hit, sample neighboring grid nodes.
- Continue until no more hits are encountered.
- Complicated calculations: use VSP to calculate the mean estimate and associated confidence interval.

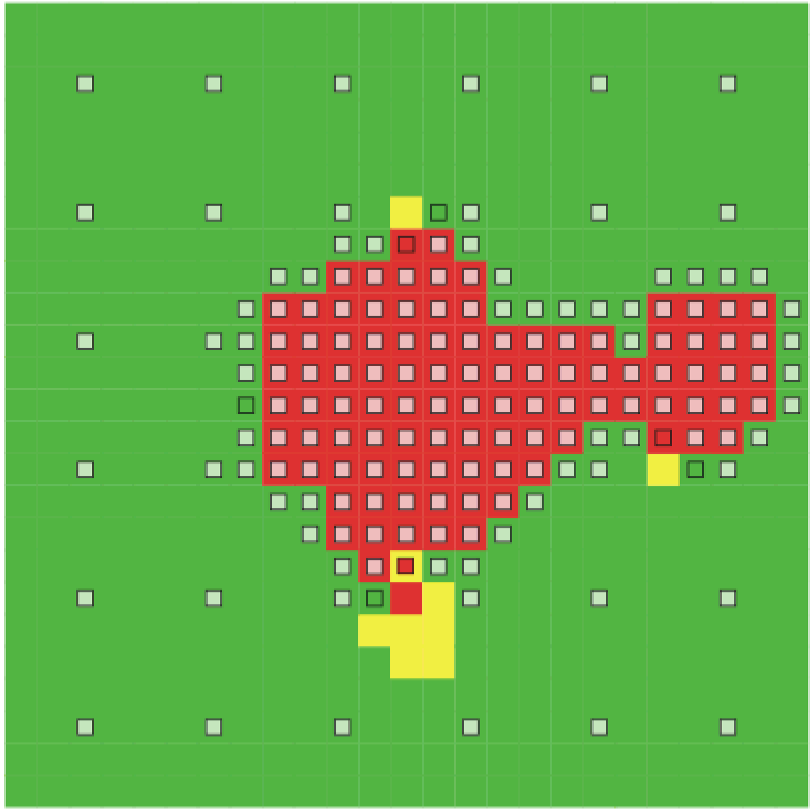
Recipe:

- Lay master grid over site.
- Start with an initial set of gridded samples, either determined by hot spot analysis or sample mean analysis.
- For every sample that is a hit, sample neighboring grid nodes.
- Continue until no more hits are encountered.
- Use VSP to determine mean estimate and associated confidence interval.

Adaptive Cluster Results

Addressing the Unknown

- 36 samples selected initially, looking for circular hot spot with radius = 50 ft.
- 7 rounds of sequential sampling: 36, 24, 40, 43, 21, 7, and 3 samples.
- Sampling complete and footprint delineated after a total of 174 samples.
- Works when contamination is "blocky."



Adaptive GeoBayesian Approaches

Addressing the Unknown

- Goal: Hot spots and boundary delineation
- Assumptions:
  - » Appropriate real-time technique is available
  - » Yes/no sample results
  - » Spatial autocorrelation is significant
  - » Desire to leverage collaborative information
- Method uses geostatistics & Bayesian analysis of collaborative data to guide sampling program & estimate probability of contamination at any location.
- Must have appropriate investigation levels for real-time technique & estimate of spatial autocorrelation range. Mean estimates can be obtained using block kriging.
- Available in BAASS software (contact Bob Johnson, ANL)

Recipe for Adaptive GeoBayesian:

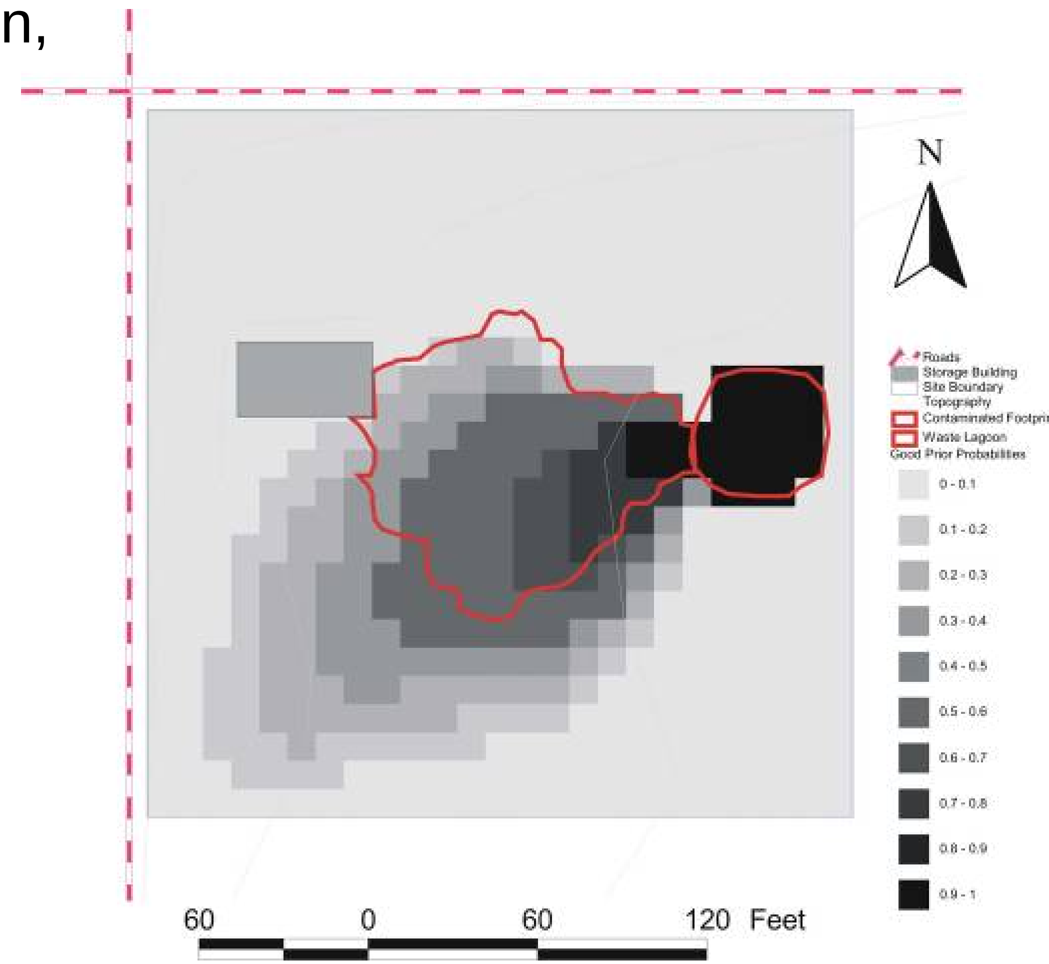
Addressing the Unknown

- Lay grid over site.
- Based on whatever information is initially available, estimate probability of contamination at each grid node.
- Convert probabilities to beta probability distribution functions.
- Specify appropriate decision-making error levels.
- Specify spatial autocorrelation range assumptions.
- Identify appropriate real-time technique and determine investigation levels.
- Implement adaptive program.

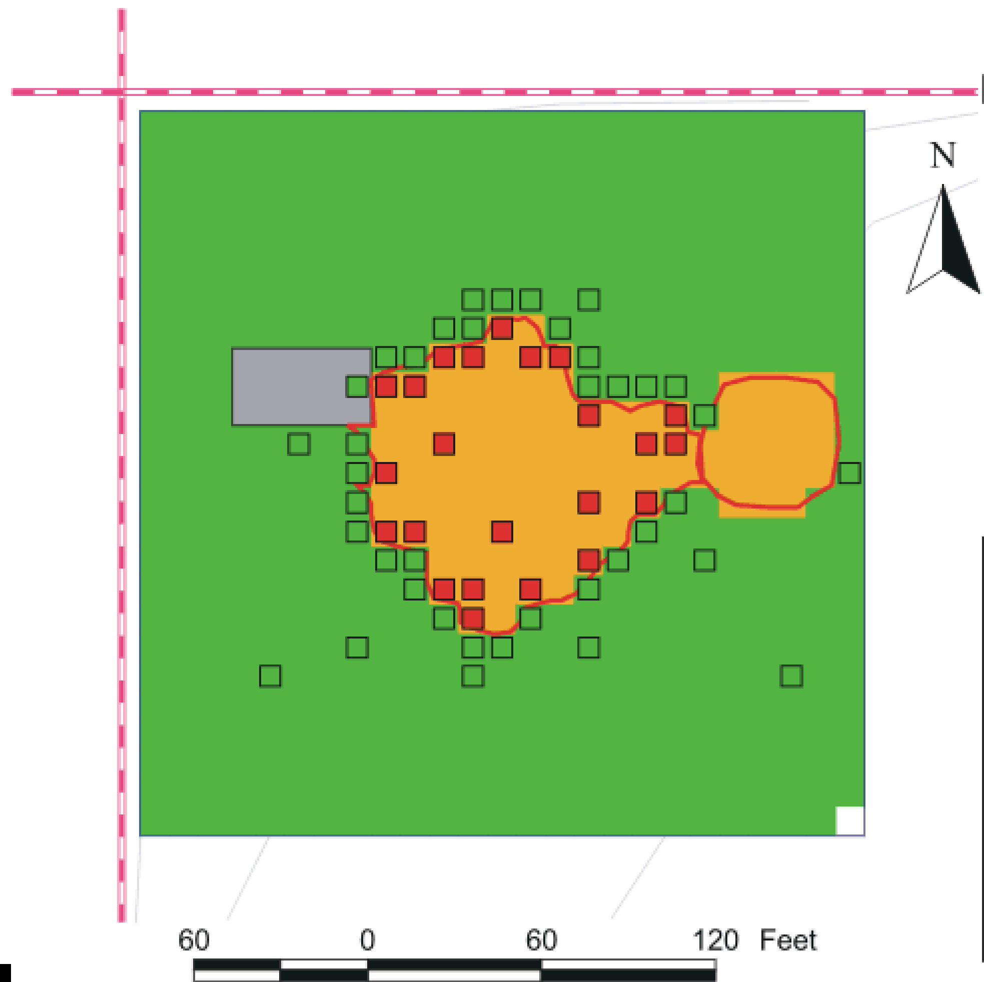
Initial Conceptual Site Model

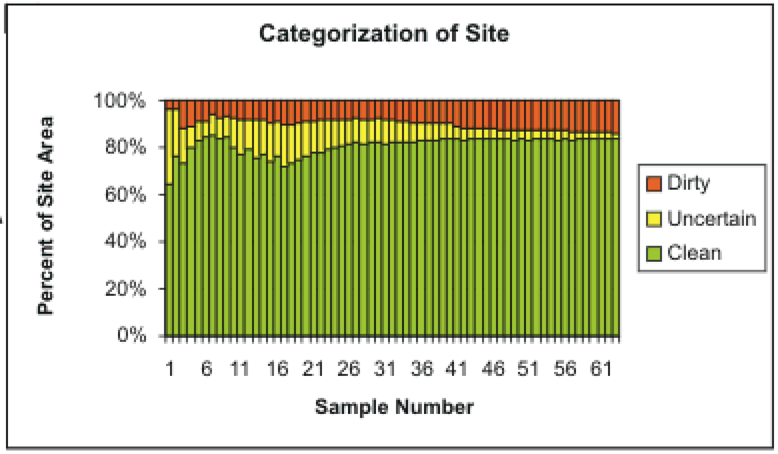
Addressing the Unknown

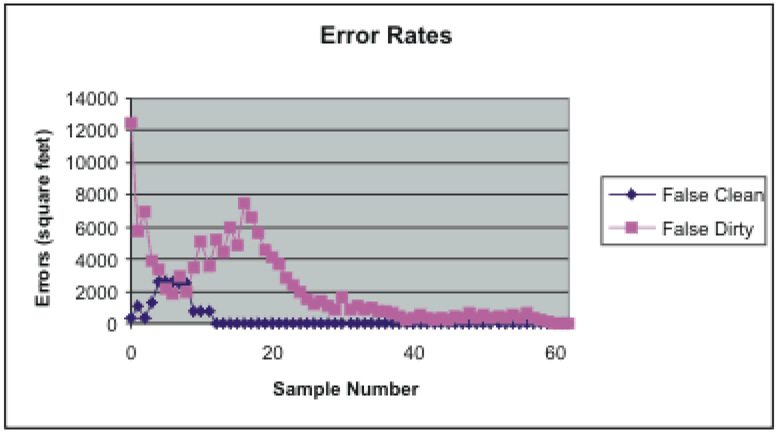
- Based on soft information, assign probability of contamination being present
- Map shows this CSM pictorially, color-coded based on contamination probability
- This CSM drives subsequent sampling decisions & becomes an important point of concurrence for stakeholders



Adaptive Sampling Program Progression...





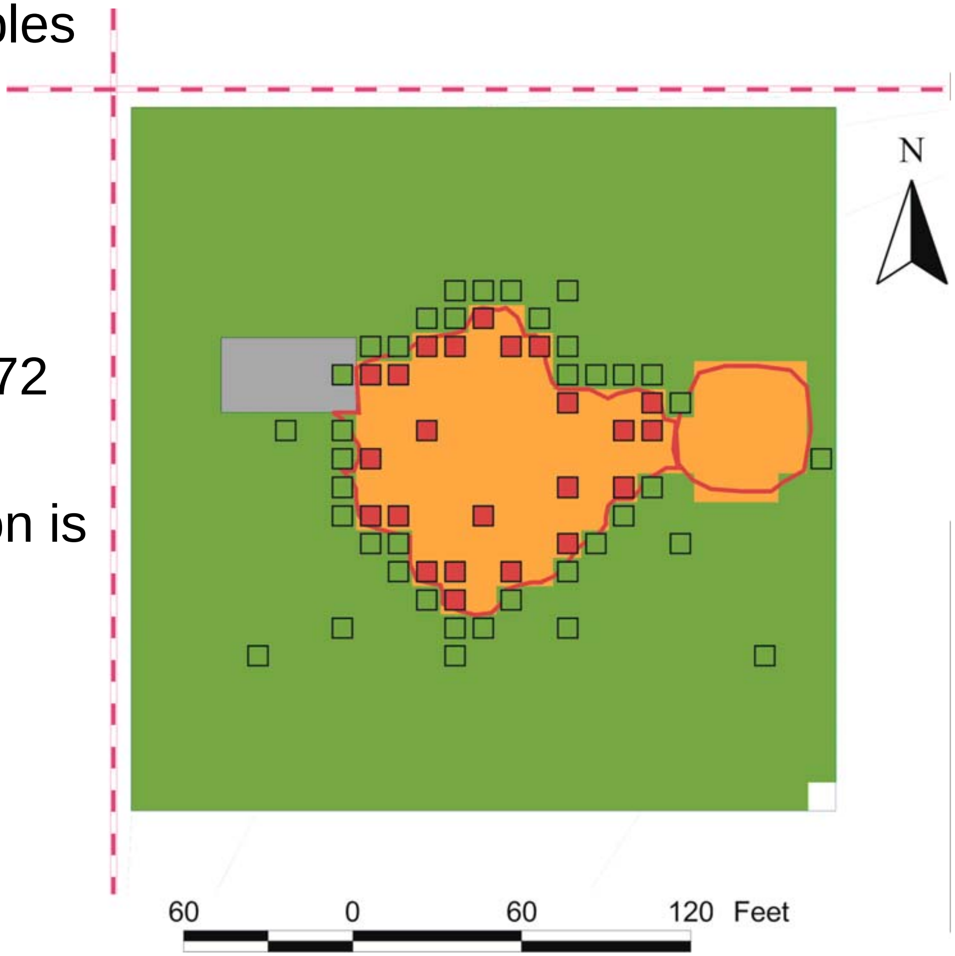


Addressing the Unknown

Adaptive Program Performance

Addressing the Unknown

- Completely done with 62 samples
- After only 22 samples, outperformed traditional 203 gridded program from an error rate perspective
- Adaptive cluster design took 172 samples
- Works best when contamination is “blocky”



The Biggest Bang Comes from Combining…

- CSM knowledge, with…
- Multi-increment sampling, with…
- Collaborative data sets, with…
- Adaptive analytics, with…
- Adaptive QC & data uncertainty reduction, with…
- Adaptive compositing, with…
- Adaptive sample location selection.