KNOWING YOUR KNOWN KNOWNS

Identifying and Reducing Data Variability & Error Through the Project Life-Cycle

Presented by:

Gavin D. Armstrong

Ohio EPA
Division of Emergency and Remedial Response

Site Investigation Field Unit

January 2011
Ohio Brownfield Conference

… in case you didn't know;

"As we know, there are known knowns. There are things we know we know. We also know there are known unknowns. That is to say we know there are some things we do not know. But there are also unknown unknowns, the ones we don't know we don't know"

- Donald Rumsfeld,
Feb. 12, 2002
Department of Defense

… so what do we know?

- A problem/condition exists that needs investigating/evaluated.
- We need to collect data to understand the potential problem/condition.
- We're going to need to have that data analyzed.

  … and
- We'll need to have those analytical results interpreted.

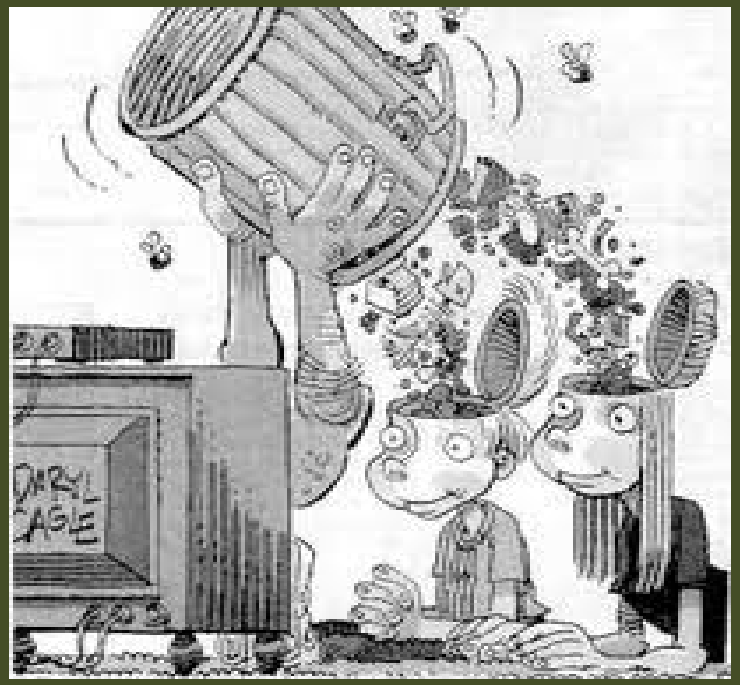
… but we also know;

$$\text{Project} = \frac{\text{Budget}}{\text{Sample cost}}$$

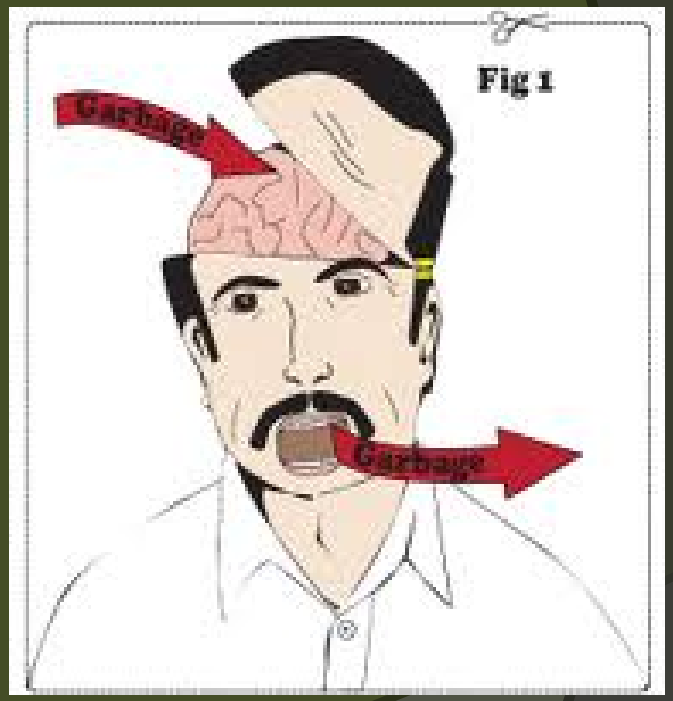
Potential error exists throughout the project lifecycle which will effect the overall quality of the final decisions being made.

Variability due to error introduces the simple concept of…

GARBAGE IN = GARBAGE OUT



=



… which can negatively impact an appropriate Risk Assessment

Sources of error;

A. Poor Planning/Sample Design
B. Field Sampling Error
C. Sample Collection/Prep Method
D. Lab Sub-sampling
E. Analytical Error

TOTAL ERROR = A + B + C + D + E

WHRE IS THE GREATEST SOURCE OF ERROR?

Sampling Error
= 90%

Lab Error
= 10%

… or for those who like to "Shout it out loud"…

Sampling Error

Lab Error

Sources of Poor Planning

- Wrong regulatory program;
  RCRA? VAP? Remedial?
  - what standards apply?
- Misunderstood end-use;
  Residential v. Commercial?
  Parking Lot v. Park?
- Population Determination;
  Decision Unit? Exposure Unit? Mixed Units?
- Wrong Analytical Method;
  Method 5035 v. 8260

Sources of Field Sampling Error

- Inadequate Numbers of Samples
- Improper Sample Locations
- Samples NOT Representative

  -Soil samples collected from a large pile

  -Vapor samples collected only once

  -Perched water zones sampled
- Lack of Consideration and/or Assumptions of Contaminant(s) of Interest

  -Vertical & Horizontal Distribution Assumptions

Improper Sample Collection/ Prep Method

- Low-flow/micro-purging v. Std. 3-Well-Volume Purging
- Homogenizing, Sieving, Dis-aggregating Soils for Metals Analysis

Error in Lab Sub-Sampling

- More specific to soil samples

  eg. 4 oz. jar of soil v. ~1 g sample for metals analysis.

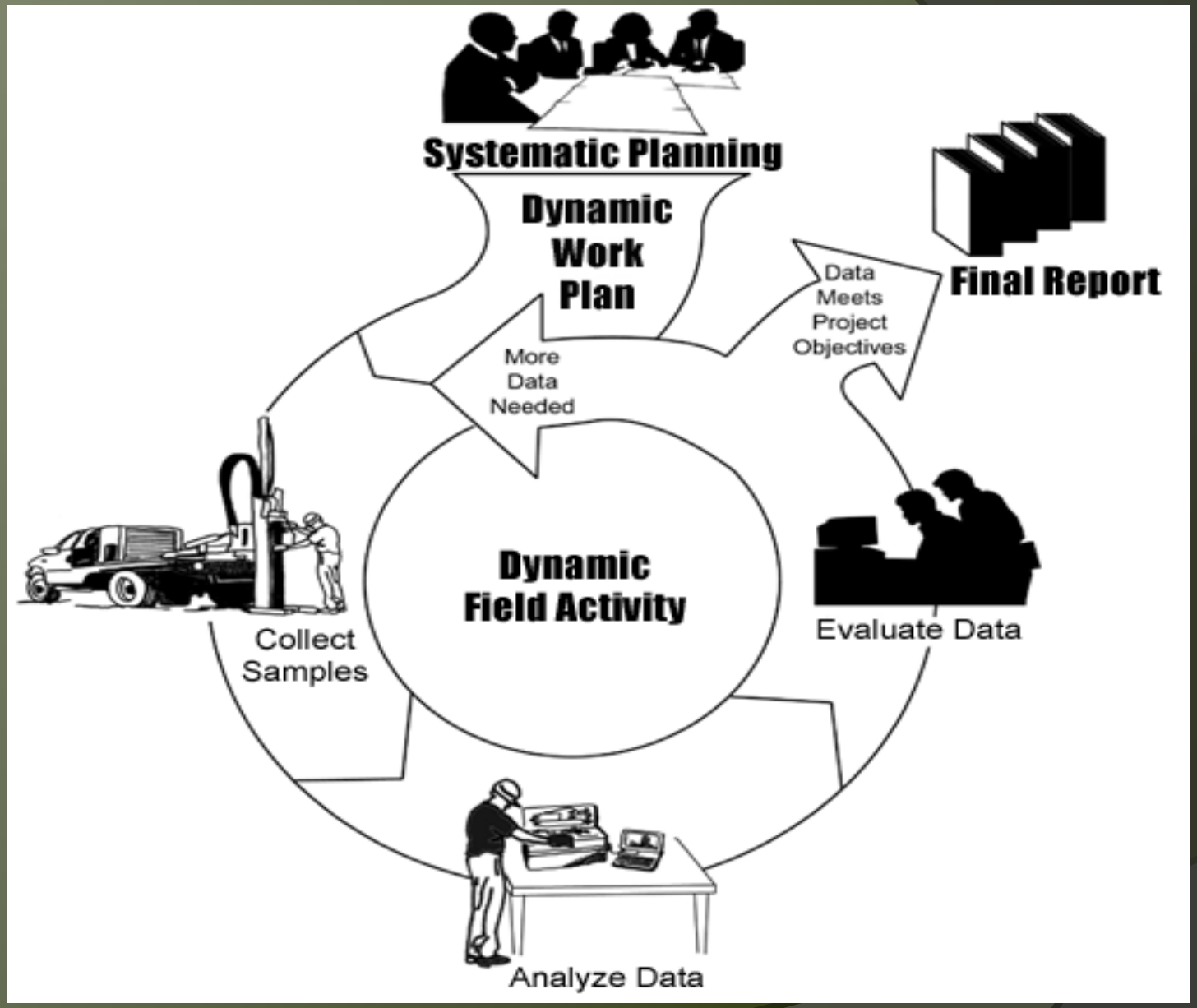
- Duplicate analyses not within acceptable range (note that some labs give themselves a wide window for sample recovery/conc.)

Sources of Analytical Error

- Sampling out of order
- Combining analytical runs
- Sample dilutions
- Data transcription/typing error
- Duplicates/reproducibility
- Inappropriate calibration
- Wrong analytical method
- Decontamination of re-usable items (cylinders)
- Sloppy/inappropriate QA/QC

Tools to Reduce Variability and/or Error

- Systematic Planning
- Dynamic Work Strategies
- On-site, Real-time Data Analysis
- Incremental Sampling Methodology
- COMMUNICATION AMONG STAKEHOLDERS



US EPA, *Using Dynamic Field Activities for On-Site Decision Making,* OSWER no. 9200. 1-40 EPA/540/R-03/002. May 2003

Conceptual Site Models & Dynamic Work Strategies

- Are the basis for all site decision-making
- Clarifies assumptions and defines the most appropriate data to gather

Preliminary CSM predicts **contaminant** distributions

Prediction guides SAP development

Data confirms or modifies predictions as CSM gradually progresses

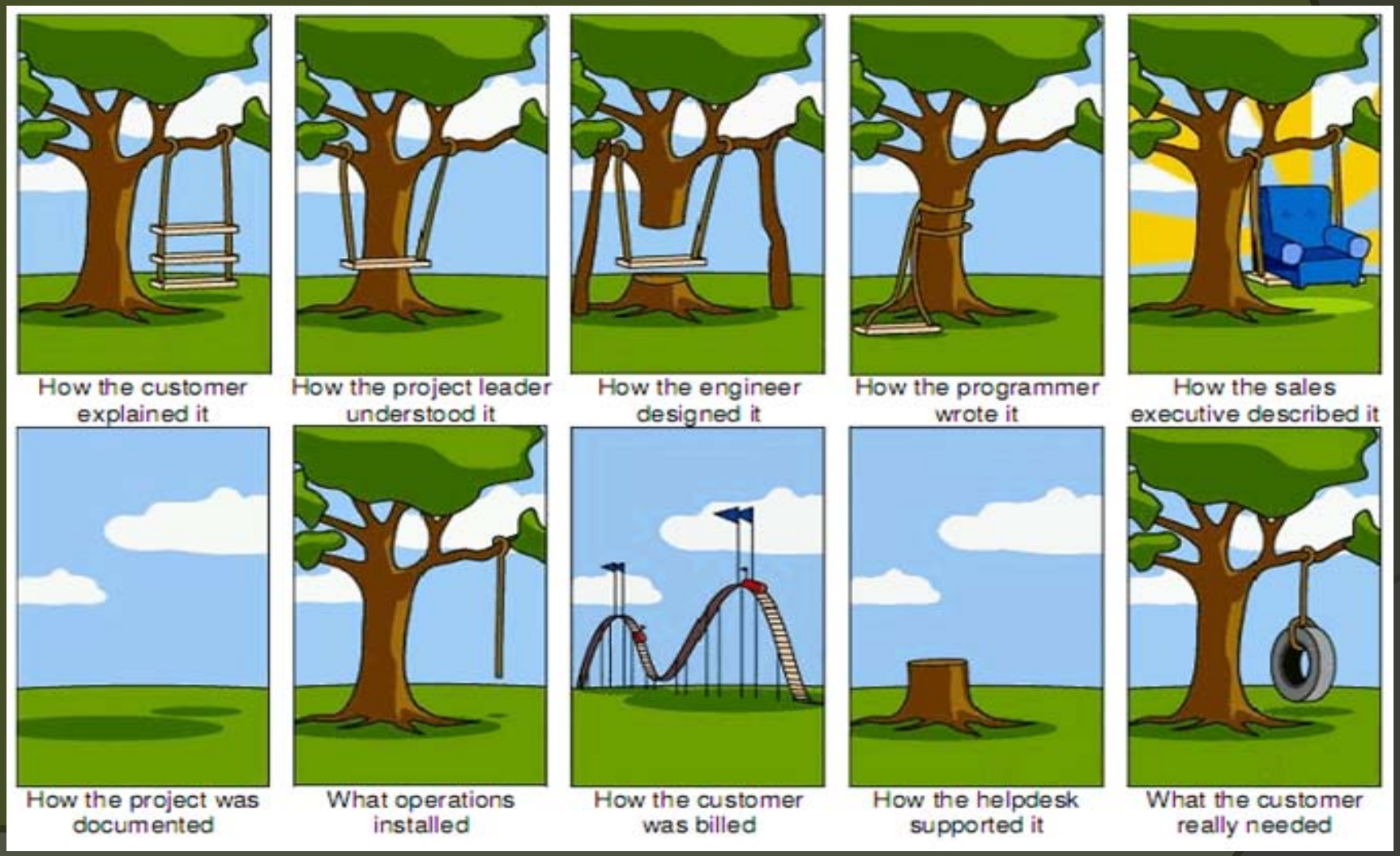
Established CSM is basis of decisions about exposure risk & risk mgt. strategies

The CSM & Dynamic Work Strategy help clarify the basic questions of….

- Where are you going to sample?
- How many samples are you going to take?
- Can you get a representative sample?
- When will I have enough data?

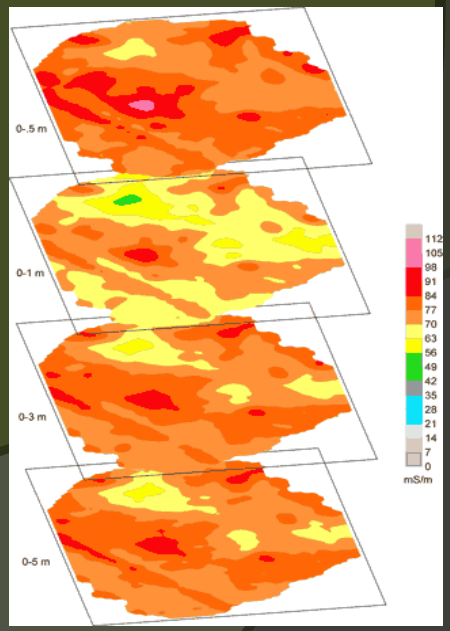
… in a "nutshell" – these are your project Data Quality Objectives

... or yet another way of looking at it...



On-site/Real-time Data

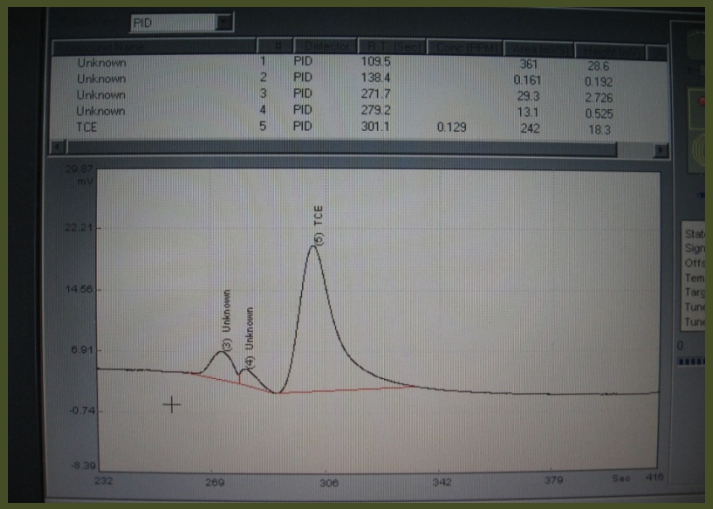
- Mobile Labs
- Portable GC, XRF, Immunoassay, colorimetric
- Field Analytical Methods
- Geophysical tools (EM, GPR, resistivity)
- Down-hole sensing tools (Membrane Interface Probe)

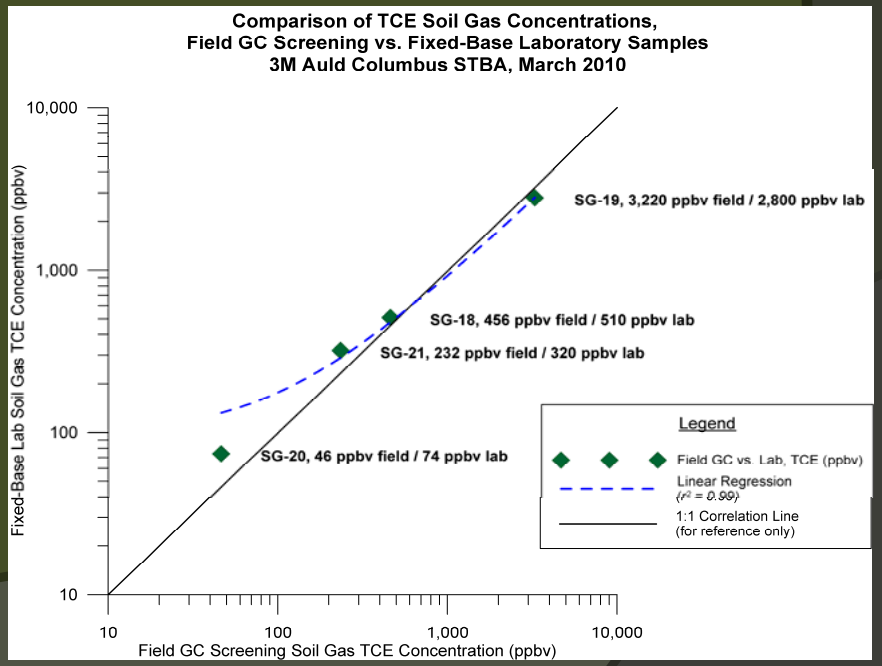


Real-Time Data Allows. . .

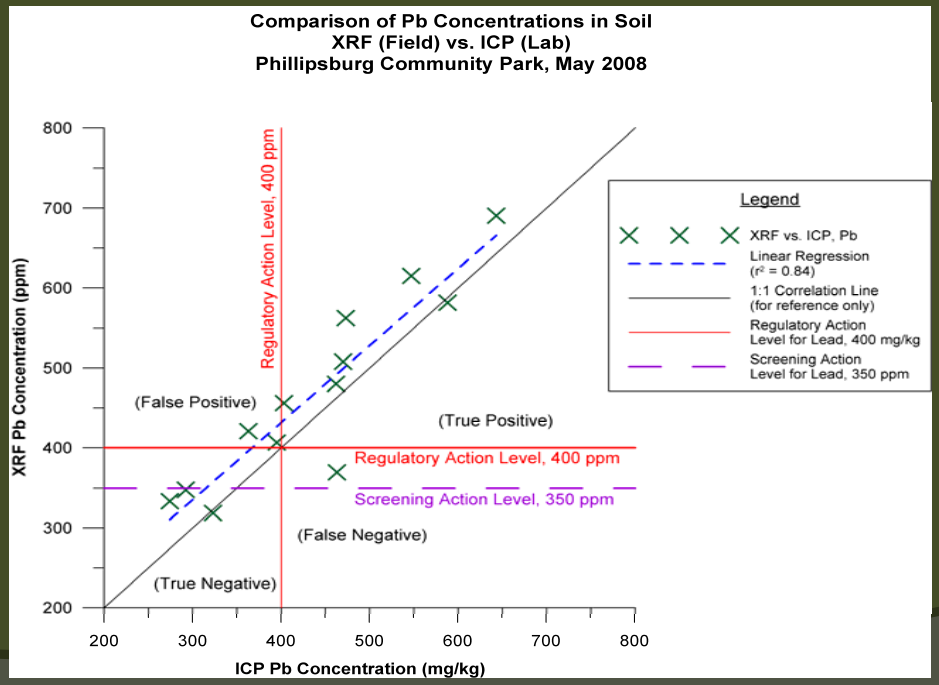
- The use of dynamic work strategies
- Continuous updates to CSM
- Target and manage greatest uncertainties as they are identified
- Optimize well placement/screen interval
- Limit mobilizations/prioritize data gaps
- Choose collaborative samples
- Combine traditionally segmented activities
- Monitor in-situ remedy effectiveness

Portable GC


PID
Unknown 1 PID 109.5 361 28.6
Unknown 2 PID 138.4 0.161 0.192
Unknown 3 PID 271.7 29.3 2.726
Unknown 4 PID 279.2 13.1 0.525
TCE 5 PID 301.1 0.129 242 18.3
(3) Unknown
(4) Unknown
(5) TCE


Comparison of TCE Soil Gas Concentrations,
Field GC Screening vs. Fixed-Base Laboratory Samples
3M Auld Columbus STBA, March 2010
Fixed-Base Lab Soil Gas TCE Concentration (ppbv)
Field GC Screening Soil Gas TCE Concentration (ppbv)
10
100
1,000
10,000
SG-19, 3,220 ppbv field / 2,800 ppbv lab
SG-18, 456 ppbv field / 510 ppbv lab
SG-21, 232 ppbv field / 320 ppbv lab
SG-20, 46 ppbv field / 74 ppbv lab
Legend
Field GC vs. Lab, TCE (ppbv)
Linear Regression
(r^2 = 0.99)
1:1 Correlation Line
(for reference only)

X-Ray Fluorescence



Immunoassay
Hach Test Kits
Spectrophotometer

Incremental Sampling…

"Incremental sampling methodology (ISM) is a structured sampling and processing protocol having specific elements designed to reduce data variability and increase sample representativeness for a specified area/volume of soil under investigation"

ITRC (Interstate Technology & Regulatory Council). 2011. *Incremental Sampling Methodology.* ISM-1. Washington, D.C.: Interstate Technology & Regulatory Council, Authorizing Team. www.itrcweb.org

ISM Provides Consideration for;

PLANNING – IMPLEMENTATION - ASSESSMENT

Reducing Fundamental Error

Collection of Truer Representative Sample

Spatial and Temporal Samples

Probabilistic Sampling

Equalizing the Mass of the Sample and Mass of the Target Population

Systematic Planning

Homogeneity NOT Heterogeneity

Randomness

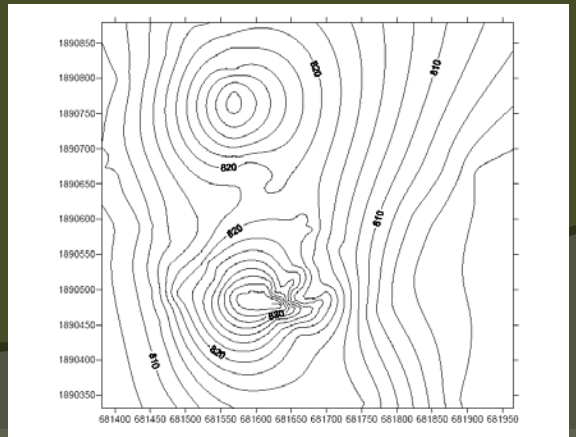
Gy's Sampling Theory

...essentially ...

Utilization of additional resources - key components to ACQUIRING BETTER DATA...

Trimle R8 GPS

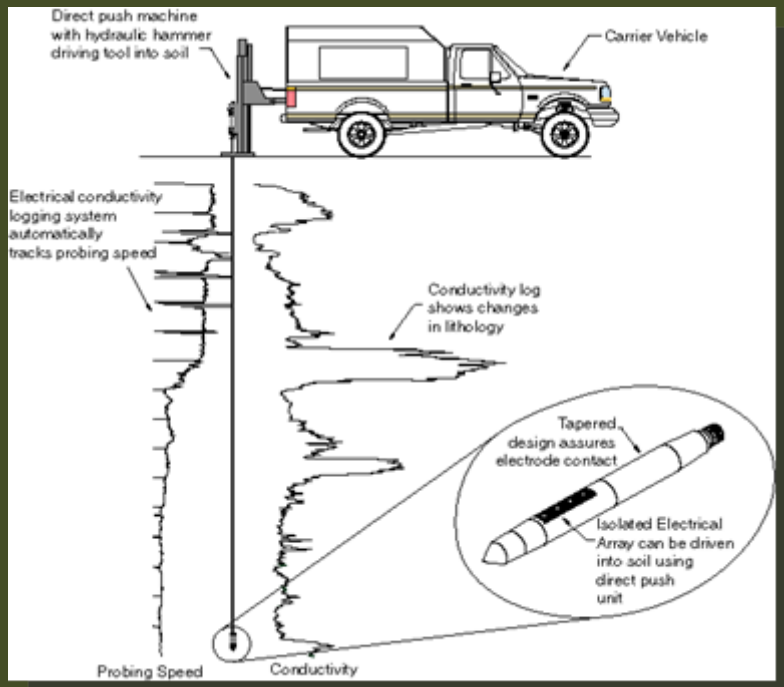
Trimble Dini Leveling Instrument

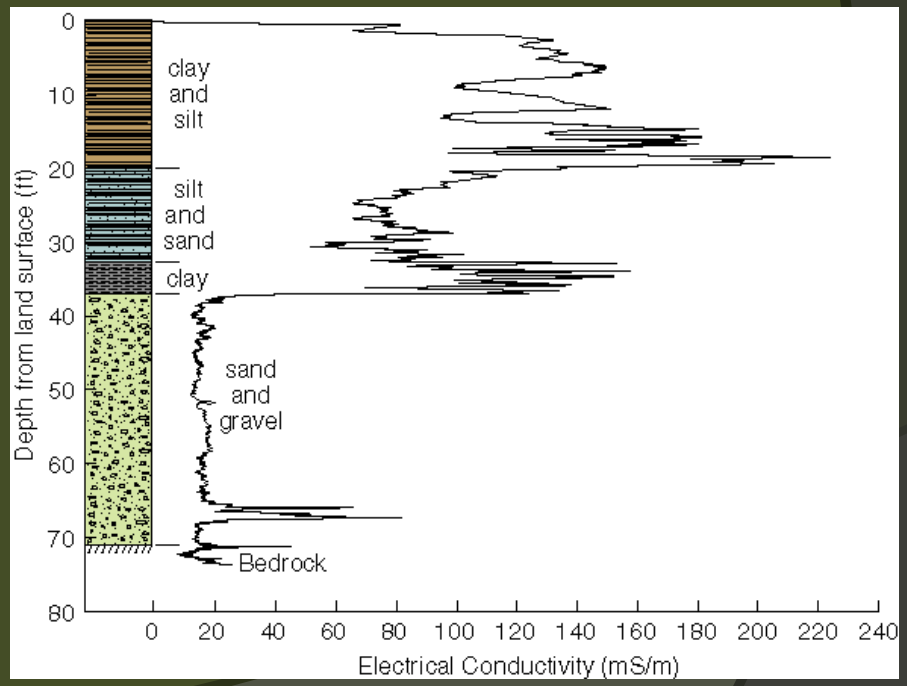


Geoprobe Capabilities…

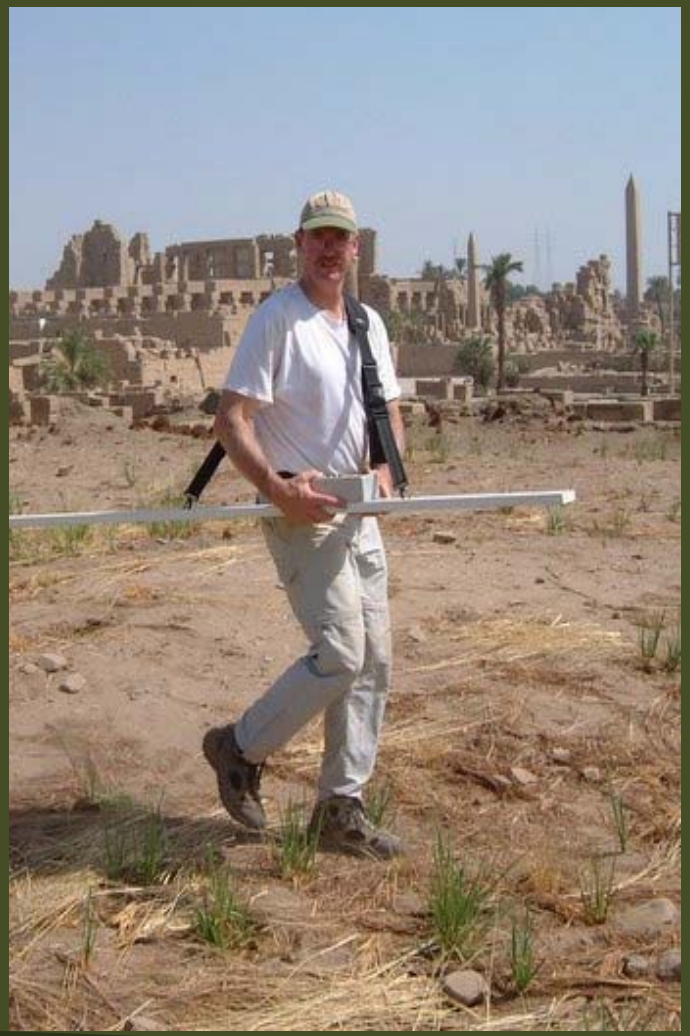
Soil Conductivity

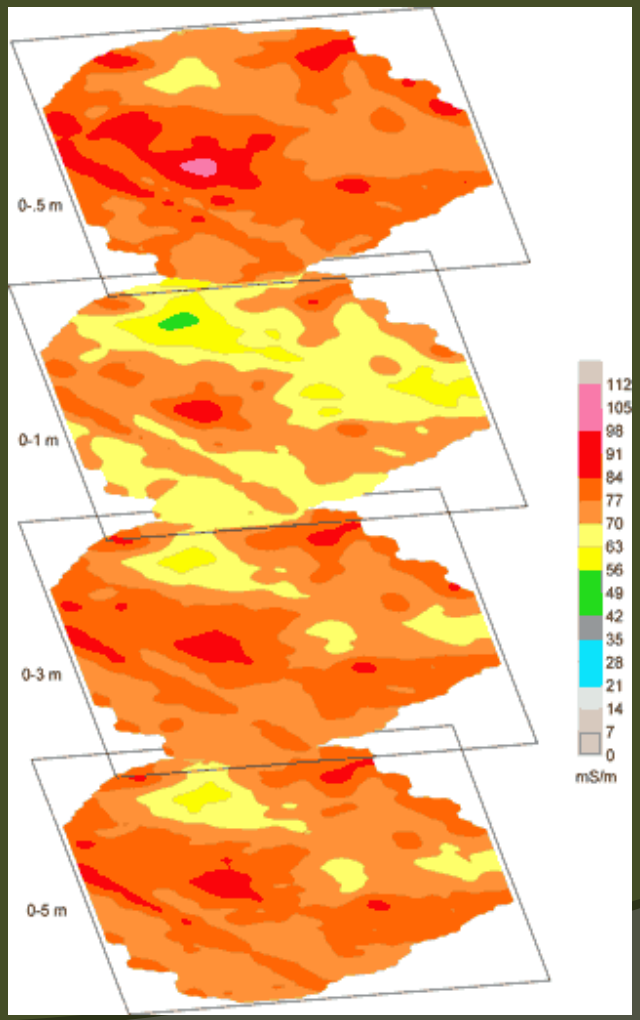
Soil Conductivity Probe - maps subsurface lithology through soil conductivity measurements; capable to ~60 feet bgs.

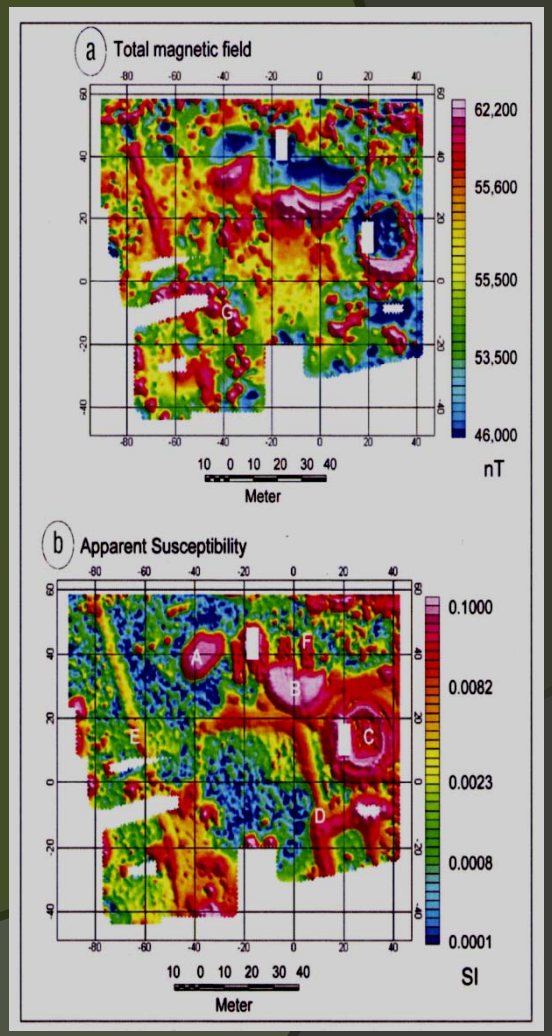




Electromagnetic Sensor / Ground Conductivity Meter







... in summary...

PLANNING: *Systematic Planning*

Data Quality Objectives

Conceptual Site Model

IMPLEMENTATION:

Real-Time Data Collection

Utilize Various Tools

On-Site Decision-Making

ASSESSMENT:

Evaluate Data

Re-Evaluate DQOs

Utilize a *Synergy of Professional Specialties* to Answer Three Basic Questions;

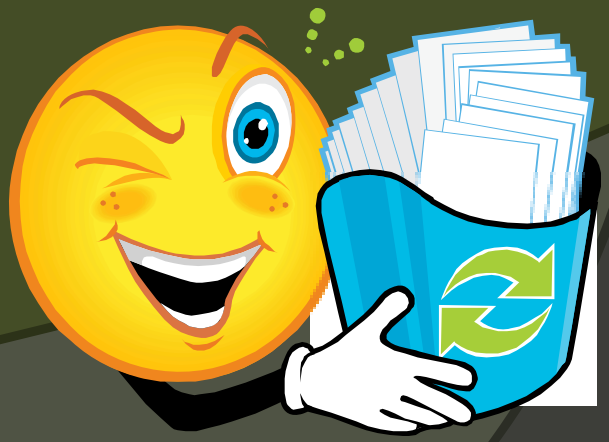
WHAT DO YOU NEED?

WHY DO YOU NEED IT?

HOW WILL YOU USE IT?

Technical Support Resources/Services

- Sampling design, scope of work, sequencing
- Create/refine Conceptual Site Model
- Assist with systematic planning
- DQO development
- Work-plan development and/or review
- Effective use of real-time measurements, field analytics, on-site analysis
- Statistical data analysis
- Decision Support Tools
- Etc., etc.



Available Resources

- Ohio EPA – DERR/VAP Staff
  www.epa.state.oh.us/derr
- USEPA Brownfields Technical Support Center
  www.brownfieldstsc.org – or - http://www.btsc.org
- Triad Resource Center
  http://www.triadcentral.org
- Hazardous Waste Cleanup Information (CLU-IN) Internet Site
  http://clu-in.org
- Field Analytical Technology Encyclopedia (FATE)
  http://fate.clu-in.org/
- Interstate Technology Regulatory Council (ITRC)
  http://www.itrcweb.org