# The Critical Role of Data Management

## ConSoil 2008 – Milan

Carlos Pachon

U.S. EPA

Technology Innovation & Field Services Division

pachon.carlos@epa.gov

703-603-9904


Robert Howe

Tetra Tech EM Inc.

robert.howe@ttemi.com


Tom Purucker

U.S. EPA

Office of Research and Development

purucker.tom@epa.gov

1

# Course Objectives

◆ Understand the lifecycle of project data

  » Collection and transfer

  » Storage and processing

  » Communicate results

  » Visualization and decision-making

◆ Gain an appreciation for electronic information collection techniques and data visualization tools

◆ Increase your knowledge of available tools, resources, services, and where to go for assistance

# Contaminated Site Cleanup in the U.S.

Five major markets

1.  Federal facilities, mainly Department of Defense and Department of Energy

2.  EPA's Superfund Program

3.  RCRA corrective action program

4.  Underground storage tanks

5.  "Brownfields" and State programs

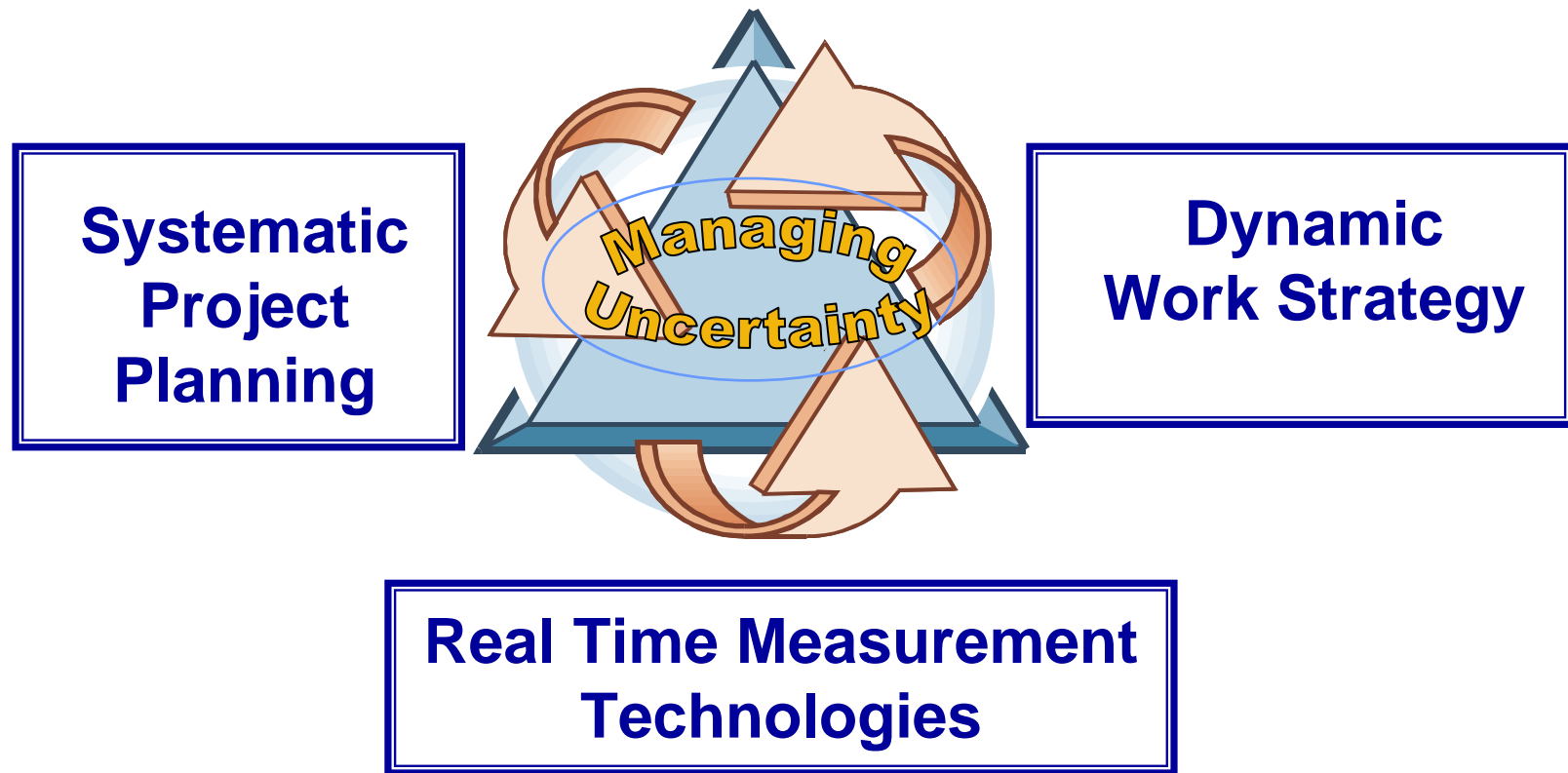# Estimated Number of Sites and Remediation Cost In the US ('04-'33)

| Program | Sites | Cleanup Cost |
|---|---|---|
| NPL | 1,146 - 1,926 | $41 - 103 B |
| RCRA, CA | 3,829 | $31 - 58 B |
| RCRA, UST | 215,827 - 395,827 | $27 - 49 B |
| DOD | 6,199 | $31 B |
| DOE | 5,000 | $73 B |
| Civilian Agencies | 3,000 | $15 - 22 B |
| States & Private | 150,000 | $ 30 B |
| Total Range<br>Middle Value | 385,001 - 565,781<br>475,000 | $248 - 366 B<br>$302 B |

# The Triad Approach
## Second Generation Practices



**Systematic Project Planning**

**Managing Uncertainty**

**Dynamic Work Strategy**

**Real Time Measurement Technologies**

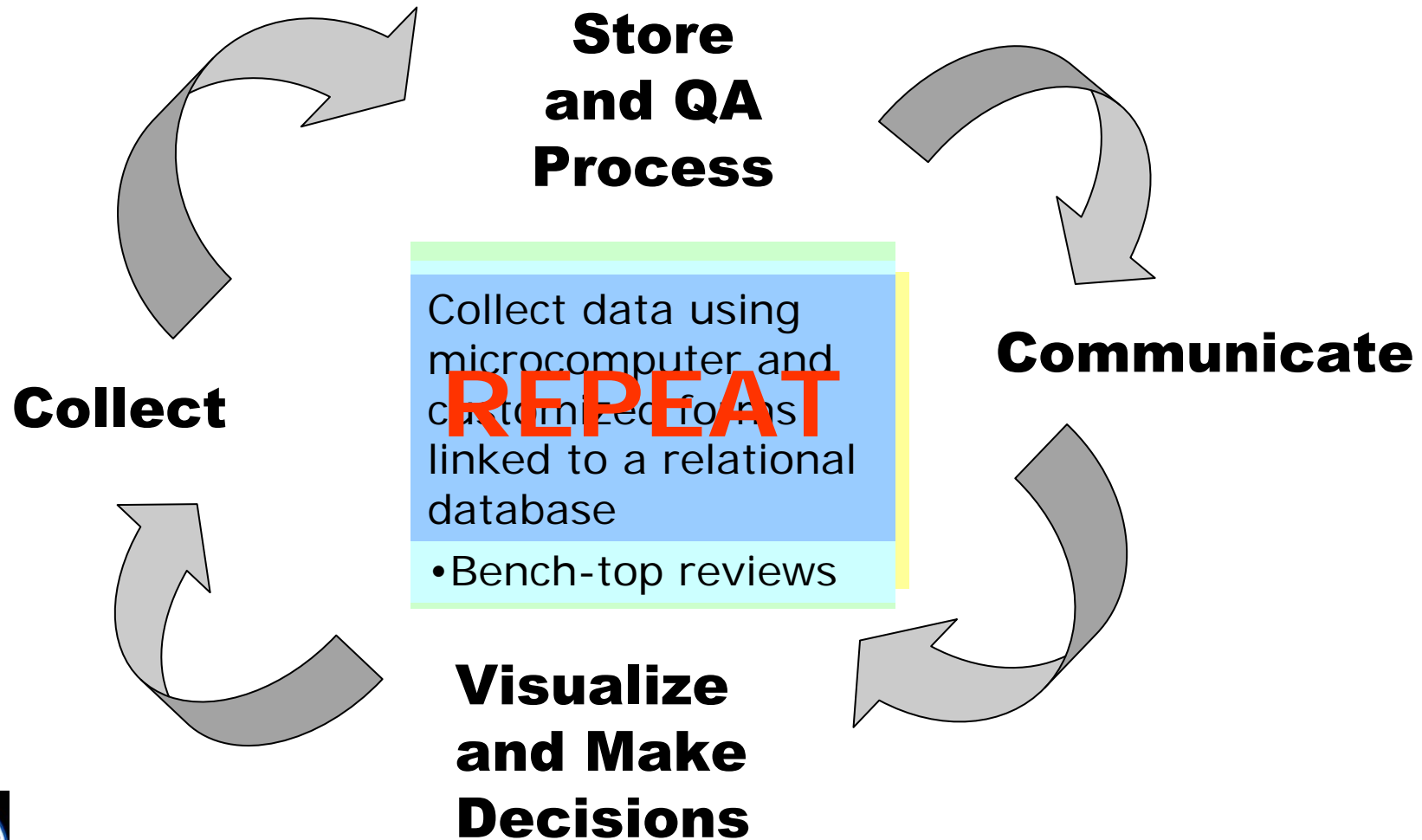*Synthesizes practitioner experience, successes, and lessons-learned into an institutional framework*

# Field sampling and analytical equipment provides more data; however data ≠ information

# A Simplified Project Data Lifecycle

**Store and QA Process**

**Collect**

**Communicate**

Collect data using microcomputer and customized forms linked to a relational database

REPEAT

• Bench-top reviews

**Visualize and Make Decisions**
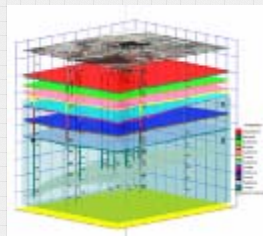
# The Big Picture: Data Flow & Tools

## QA/QC

**Collect Data**

Field tools (eg XRF)

Data mgmt tools (ex, EDD, Scriblets, Forms II Lite, etc. at EPA)

Field Database (Scribe)

Regional Data Repository (WQX/STORET, EQuIS)

*Store Data*

*Process Data*

Field Data

Lab Data

Database

Evolving Conceptual Site Model

*Make Decisions*

*Communicate*

Scribe.net

EPA OSC Website

Quickplace

Collaboration Pages
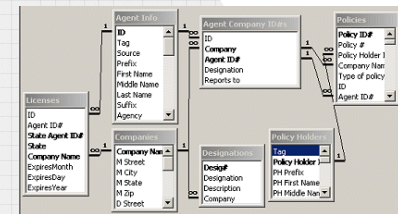
Web Conferencing

MAROS

F/S Plus

FIELDS Tools

VSP

SADA

DST Matrix

EVS

Distance Collaboration

Decision Support Tools

Data Visualization Tools

# Collect Data



**Field Data**

**Lab Data**

**From Field tools:**

**(eg XRF, LIF, EC, CPT, MIP, FFD, Mobil Lab)**

**To:**

**Scriblets**

**Forms II Lite**

**R5 EDD,SEDD**

◆ Continuing advances in technologies generating real time field data

◆ Data can be directly downloaded into electronic data deliverables (EDD) to deliver preliminary results to a relational database

&raquo; Examples:  XRF, LIF, FFD, EC, CPT, MIP, Immunoassays, Mobil Lab

◆ Flexible database formats vs. standardization

◆ Ease of data - Pull-down screens and customizable database forms prompt sampling crews to include metadata (Example:  Scriblets, FAST)
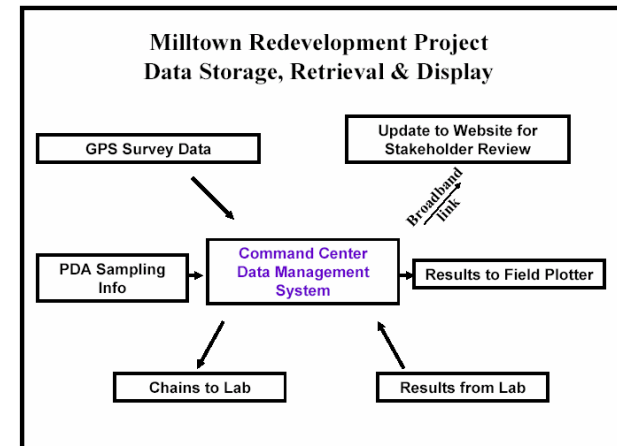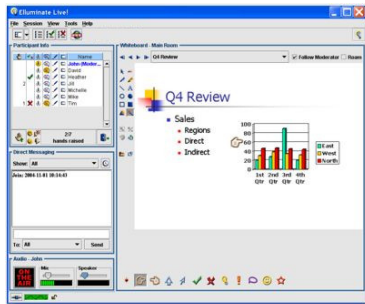
# Store and Process Data
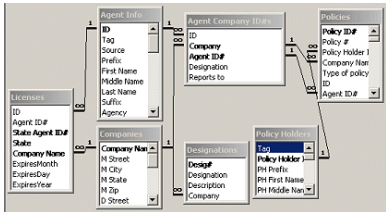
**Database**

**Field Database (Scribe)**

**Regional Data Repository (WQX/STORET, EQuIS)**

◆ Relational databases are modified to test for QA/QC requirements

◆ Data is queried, allowing site personnel to perform desktop reviews

◆ Reconciliation of identified QA/QC issues in the field

◆ Output data to applications

**Milltown Redevelopment Project Data Storage, Retrieval & Display**

GPS Survey Data

Update to Website for Stakeholder Review

Broadband link

PDA Sampling Info

Command Center Data Management System

Results to Field Plotter

Chains to Lab

Results from Lab

# Communicate Information



**Distance Collaboration**

EPA OSC Website

Lotus QuickPlace

EPA Scribe.net

Web Conferencing

**Decision makers and stakeholders are dispersed – How do we communicate with them?**

◆ Use of project/site-oriented websites where team members can quickly and securely share information

» Example:  EPA WebOSC sites

» Example:  QuickPlace sites

◆ Live virtual meeting tools allow teams to review information as if they were in the same physical space

» Example:  Go2Meeting (commercial)

» Example:  MeetingPlace (commercial)

# Make Decisions



**Decision Support Tools**

**Data Visualization Tools**

NARAC/IMAAC

F/S Plus

FIELDS Tools

VSP

SADA

DST Matrix

EVS



◆ Field data and lab data on a common platform allow teams to use decision support tools to help with dynamic sampling designs and visualization of available information

» Examples: Visualization of site conditions at Milltown redevelopment site; Fort Lewis firing range; Lone Butte Industrial Park chlorinated solvents site
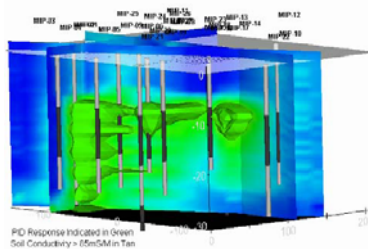
# Benefits of Real Time Data Management

◆ Increased traceability of data (defensibility)

◆ More tools to make it easier to understand results and make decisions (reliability)

◆ Improved communications internally (dependability)

◆ More confidence in decision-making (Representativeness)

# Benefits of Real Time Data Management

◆ Optimization of sampling plans to cut cost (defensibility)

◆ Stakeholder buy-in during decision makings (trust)

◆ Consensus vision on issues (efficiency)

◆ Rapid reuse (client satisfaction)

# Site Information Management:
## Data Collection, Storage, and Processing

# ConSoil 2008 – Milan

Robert A. Howe

Tetra Tech EMI

robert.howe@ttemi.com

303-441-7911

# Innovative Tools

◆ Rapid Assessment Tools (RAT) – Freeware

◆ Scriblets/Scribe – Freeware

◆ Field Assessment Survey Tool (FAST) – Freeware to EPA Region 7 and ERT

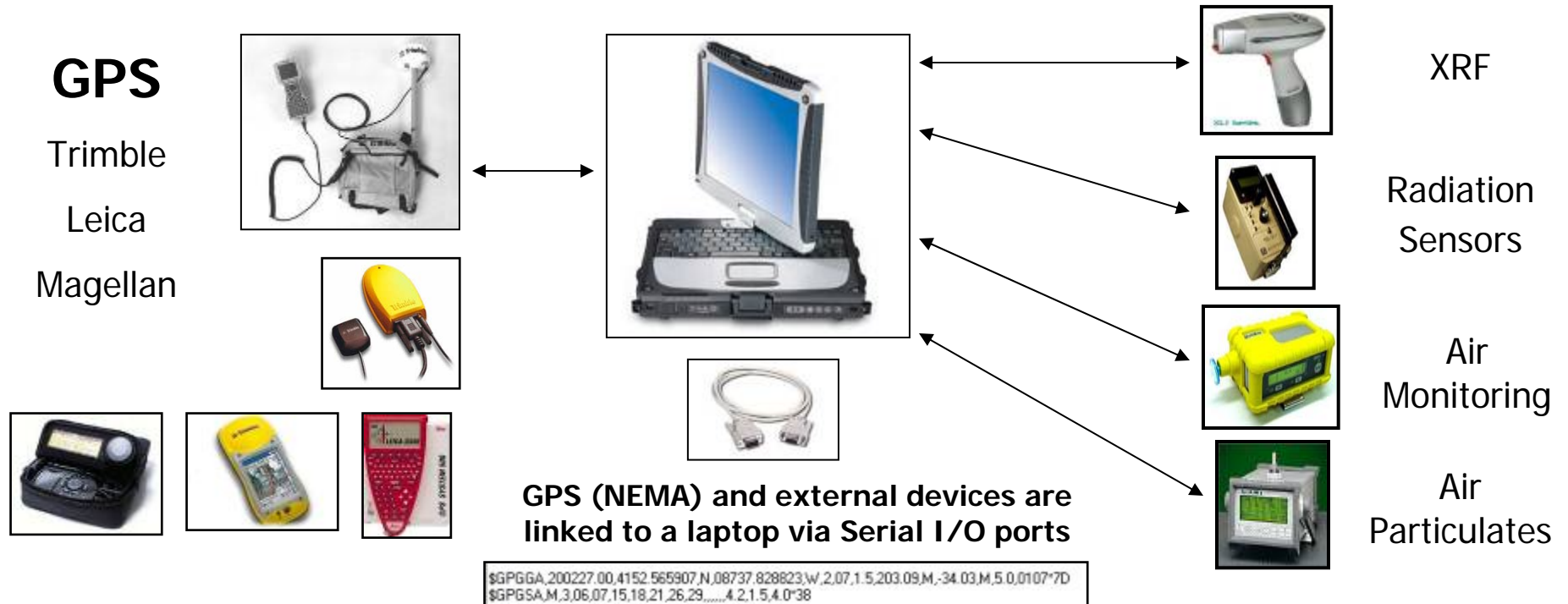# The Rapid Assessment Tools (RAT)

# RAT Architecture

Visual Basic

Conversion to .NET framework *(June, 2006)*

- ✓ multi-threading
- ✓ multi-com port retrieval

- **Geodatabase** (*March 2008*)



18

# Hardware Integration

**GPS**

Trimble

Leica

Magellan



GPS (NEMA) and external devices are linked to a laptop via Serial I/O ports

$GPGGA,200227.00,4152.565907,N,08737.828823,W,2,07,1.5,203.09,M,-34.03,M,5,0,0107*7D
$GPGSA,M,3,06,07,15,18,21,26,29,.....,4,2,1.5,4.0*38

XRF

Radiation Sensors

Air Monitoring

Air Particulates
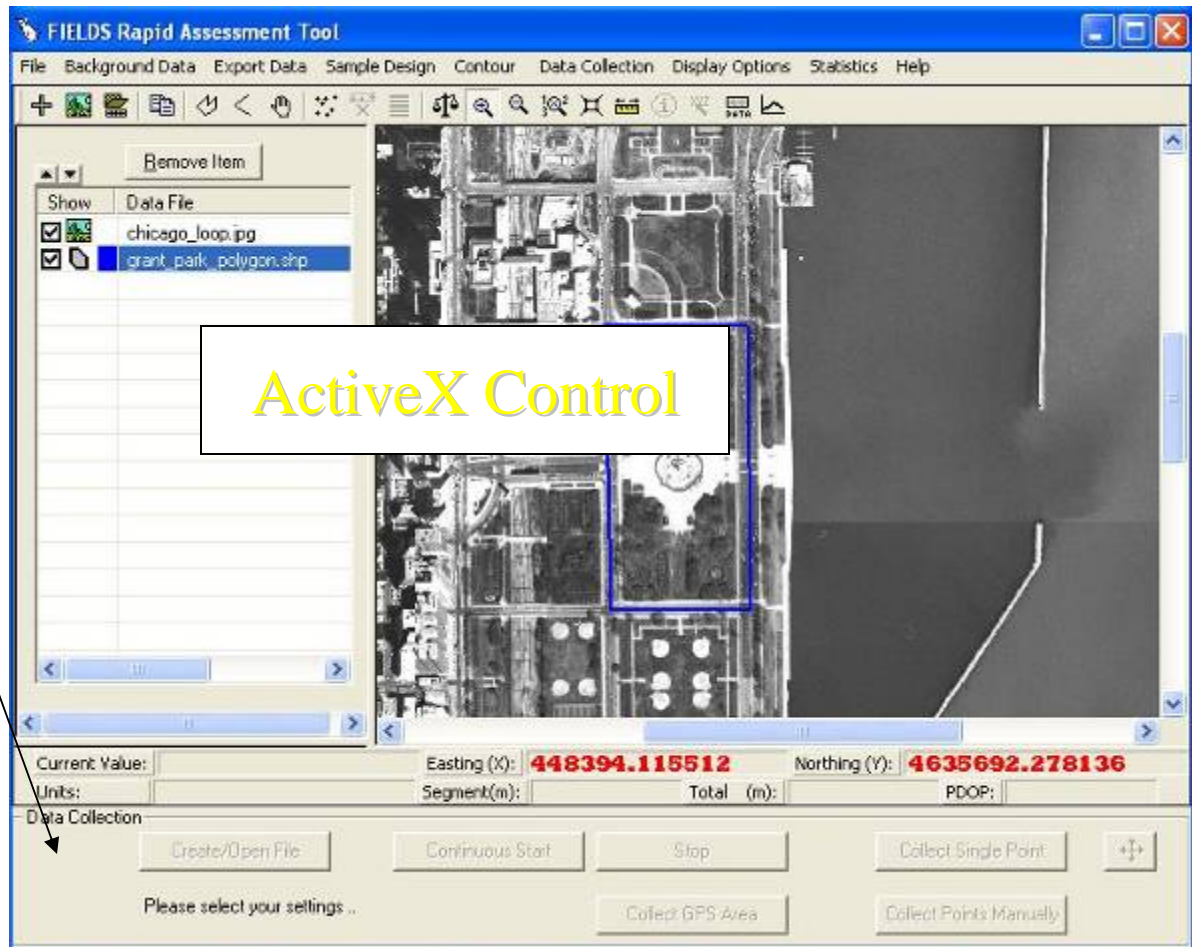
◆ Any GPS that sends a standard NMEA string can be used depending on accuracy required. This gives you real-time locations requiring no post processing.

◆ Any sensor device both analog and digital can be incorporated into the software. Up to three sensors with GPS can be configured at one time.

# Integrated Devices

- ◆ Radiation Devices
  - » Ludlum 2221 Data logger
    - — single value data output
    - — half-duplex (data out only)
  - » Ludlum 2350 & 2241 Data Logger
    - — Full Duplex (logger sends & receives)
    - — Scaler, count, & dose dump)
  - » Inovision 451P
- ◆ Air Monitoring Devices
  - » MultiRAE
    - — Over 14 sensors available from RAE Systems (CO, $H_2S$, VOCs, $O_2$, LEL, etc.)
  - » Data RAM
    - — Particulates, Temperature, Humidity, etc.
  - » Draeger Multiwarn
    - — (CO, $H_2S$, VOCs, Toxics)
- ◆ Soil Monitoring Devices
  - » XRF
    - — Over 23 different Metals
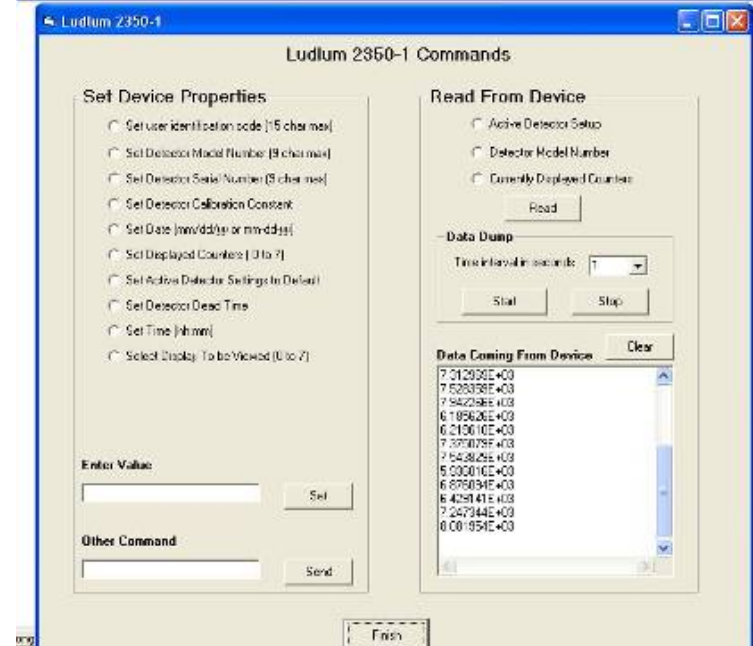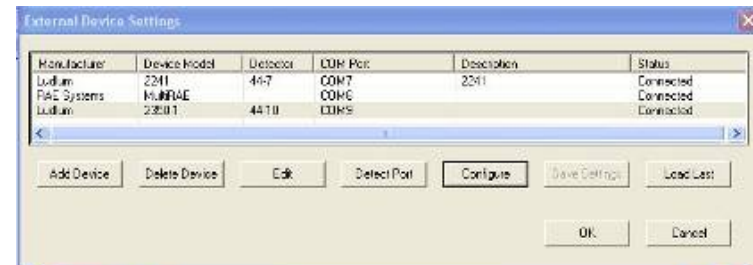
# RAT Data Storage and Processing

Import Options

◆ Background images, can be imported into projects for better visualization of the study site, defining boundaries, or creating a sample design

◆ AutoCAD files and ESRI Shapefiles can be loaded into projects

Export Options

◆ RAT data can be exported to the EPA Scribe database http://www.ertsupport.org/, or Microsoft Excel or Access

◆ Spatial image index querying capabilities allow users to obtain geographic data for projects using a lookup table method

# RAT

Advantages:

◆ Data directly downloaded from instruments

◆ Locations tied to GPS Locations

◆ Data storage in data base compatible with visualization applications

Limitations:

◆ Not compatible with MIP, LIF, or other real time tools

◆ Visualization programs (Fields) are slow

# FAST: PDA Field Application Concept



**FAST**

**On-Site Data Collector**

**Command Center/Base**

**Analytical Laboratories**

**Office Use**

**Web Network (Web Database)**

**Server & Database Maintenance**

Command Center:

Print Field Sheets
Print Maps & Photos
Access to Forms II Lite/Scribe

PDA Data Collector:

Sample Data
GPS Data
Photographs
Maps

23
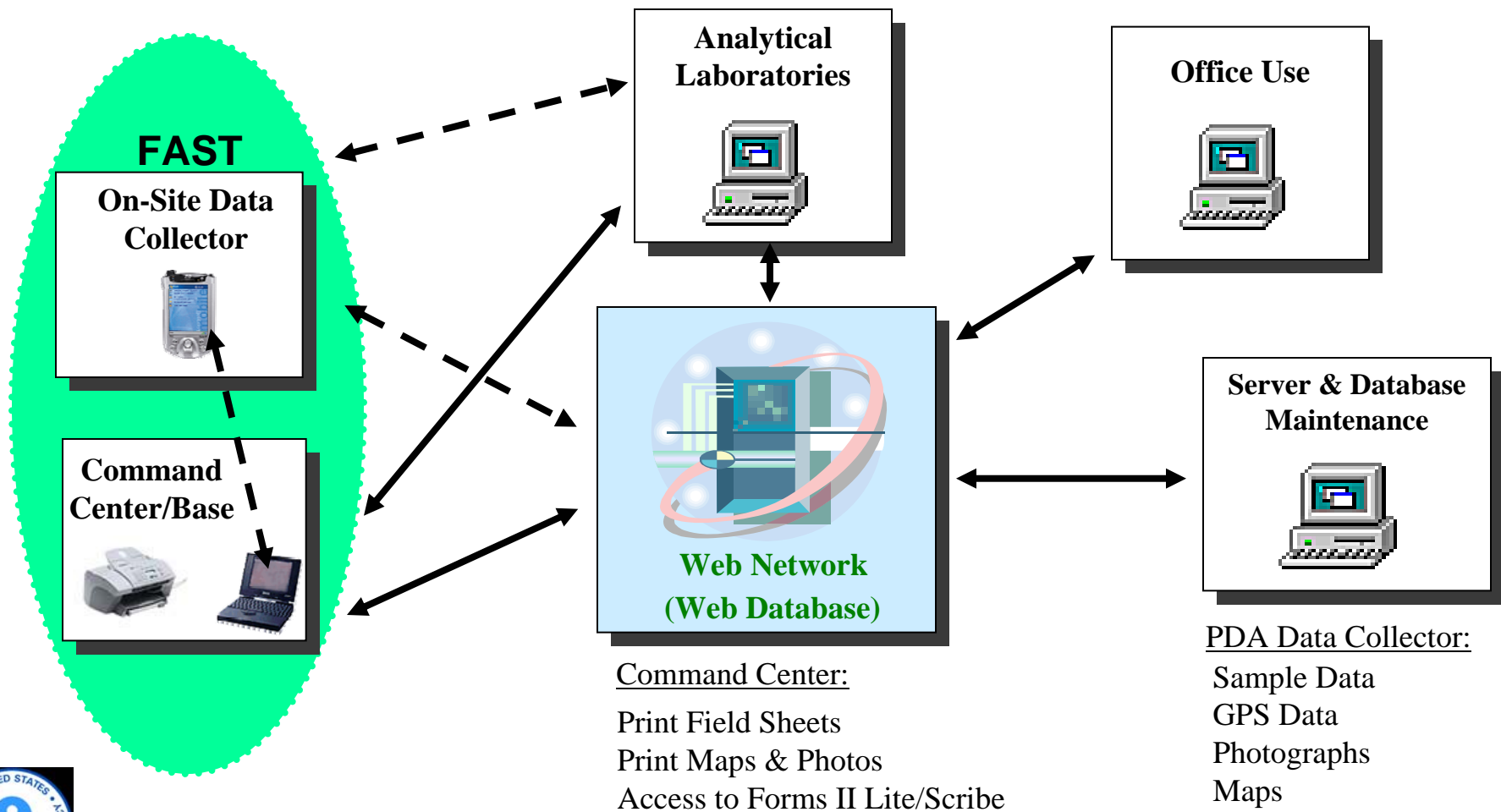
# FAST – Field Assessment Survey Tool

Advantages

◆ Custom data collection forms without programming

◆ 'Drop down menus' and 'input validation'

◆ Data can be uploaded to the relational database and/or FAST web console

Disadvantages

◆ Underlying platform outdated

◆ Cumbersome PDAs, trend is in favor of Tablet PCs

# SCRIBE: Environmental Data Management System Overview

Satellite

GPS Data

Lat/Long

Data Logger

Complete Field Data

PDA Import

Handheld/PDA

Monitoring Data

Sampling & Monitoring

Physical Samples

Result Data (EDD)

Analytical Labs (QA/QC)

database

## Data Applications

SCRIBE Data → Models — Network Station

SCRIBE Data →

SCRIBE Data → SAS — Network Station

SCRIBE Data → GIS — Network Station

SCRIBE Data → Word/WordPerfect — Network Station

25

# Scriblets – Field PDA

Advantages:

◆ Pre-determined data entry fields

◆ Pull-down menus

◆ Chain of custody and label generation

Disadvantages:

◆ Learning curve

◆ Power limitations

◆ No mapping capabilities

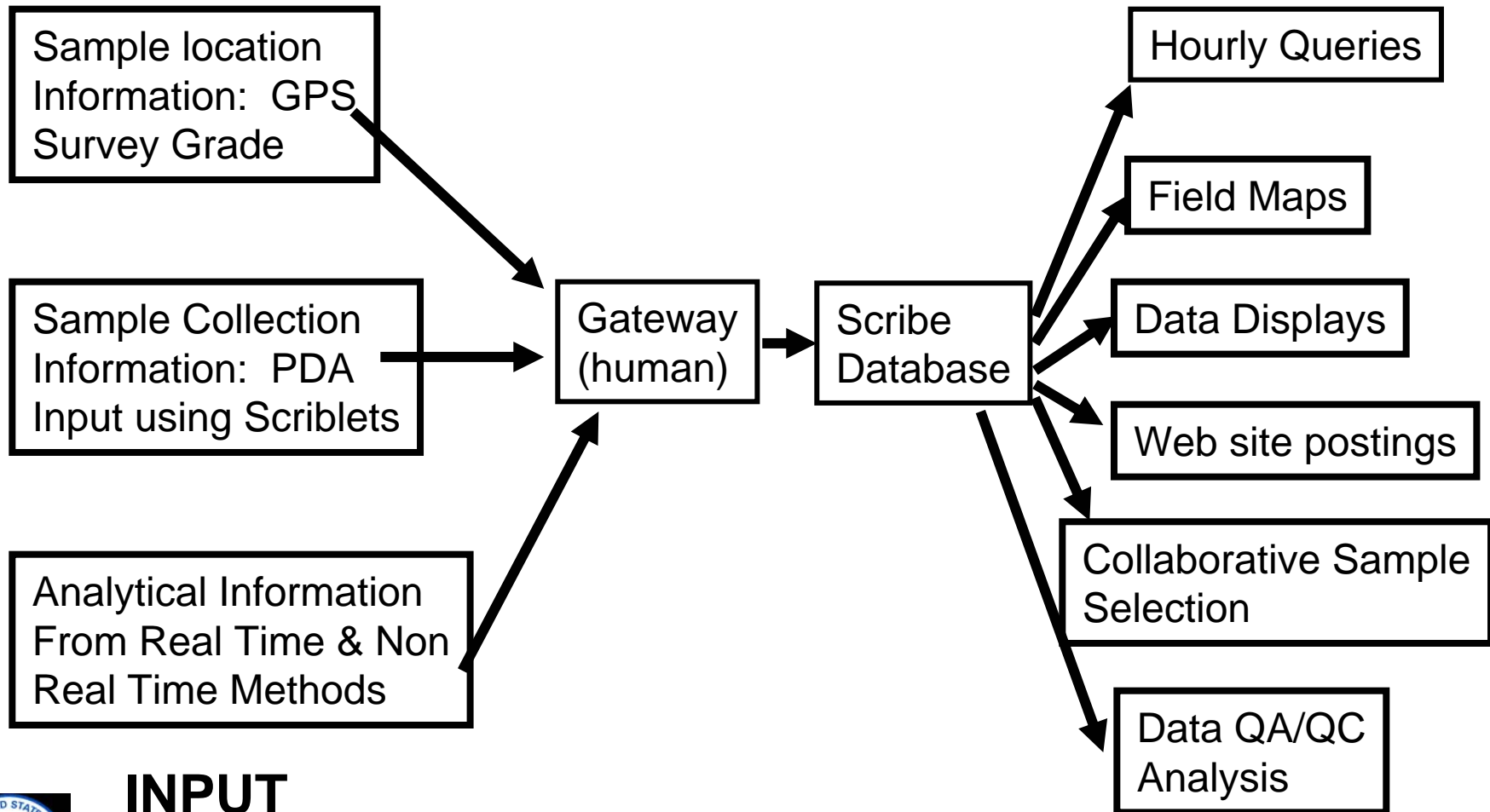# Milltown Redevelopment Project – Case Study

◆ 22 acre Brownfield site located in New Jersey

◆ Long history of industrialization but no clear past use records

◆ Slated for mixed light industrial reuse

◆ Variety of potential contaminants of potential concern

  » VOCs, TPH, metals, PAHs, PCBs

# Milltown Data Collection Logic Diagram

| | | |
|---|---|---|
| **Sample location Information:  GPS Survey Grade** | | **Hourly Queries** |
| **Sample Collection Information:  PDA Input using Scriblets** | **Gateway (human)** → **Scribe Database** | **Field Maps** |
| | | **Data Displays** |
| | | **Web site postings** |
| **Analytical Information From Real Time & Non Real Time Methods** | | **Collaborative Sample Selection** |
| | | **Data QA/QC Analysis** |

**INPUT**

**OUTPUT**

28

# Survey Grade GPS Based Instruments Used to Provide X,Y,Z Coordinates

# PDA Link to Laptop with Scribe Software

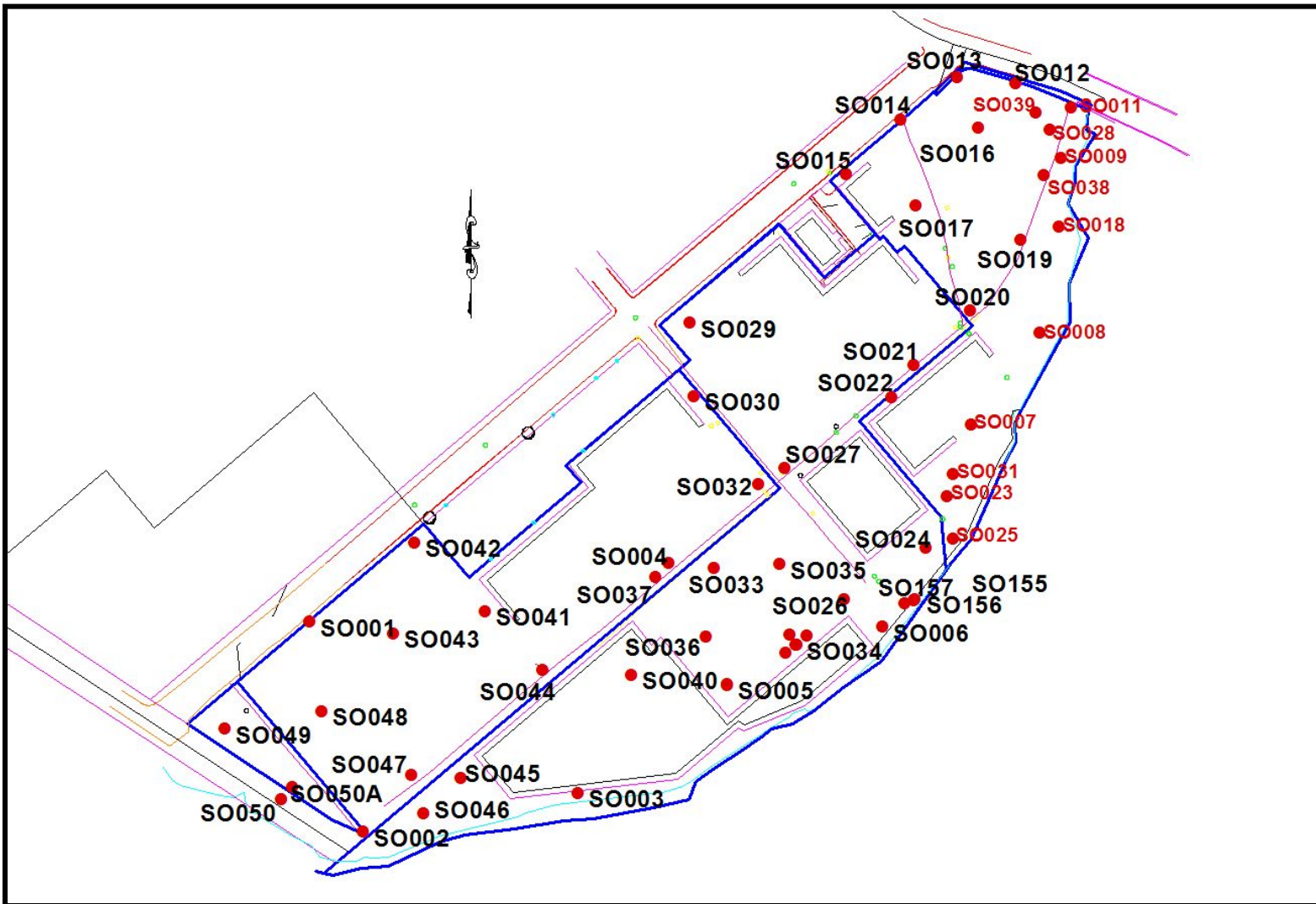# Three Mobile Laboratories Used to Generate Analytical Data

# Milltown Redevelopment Project
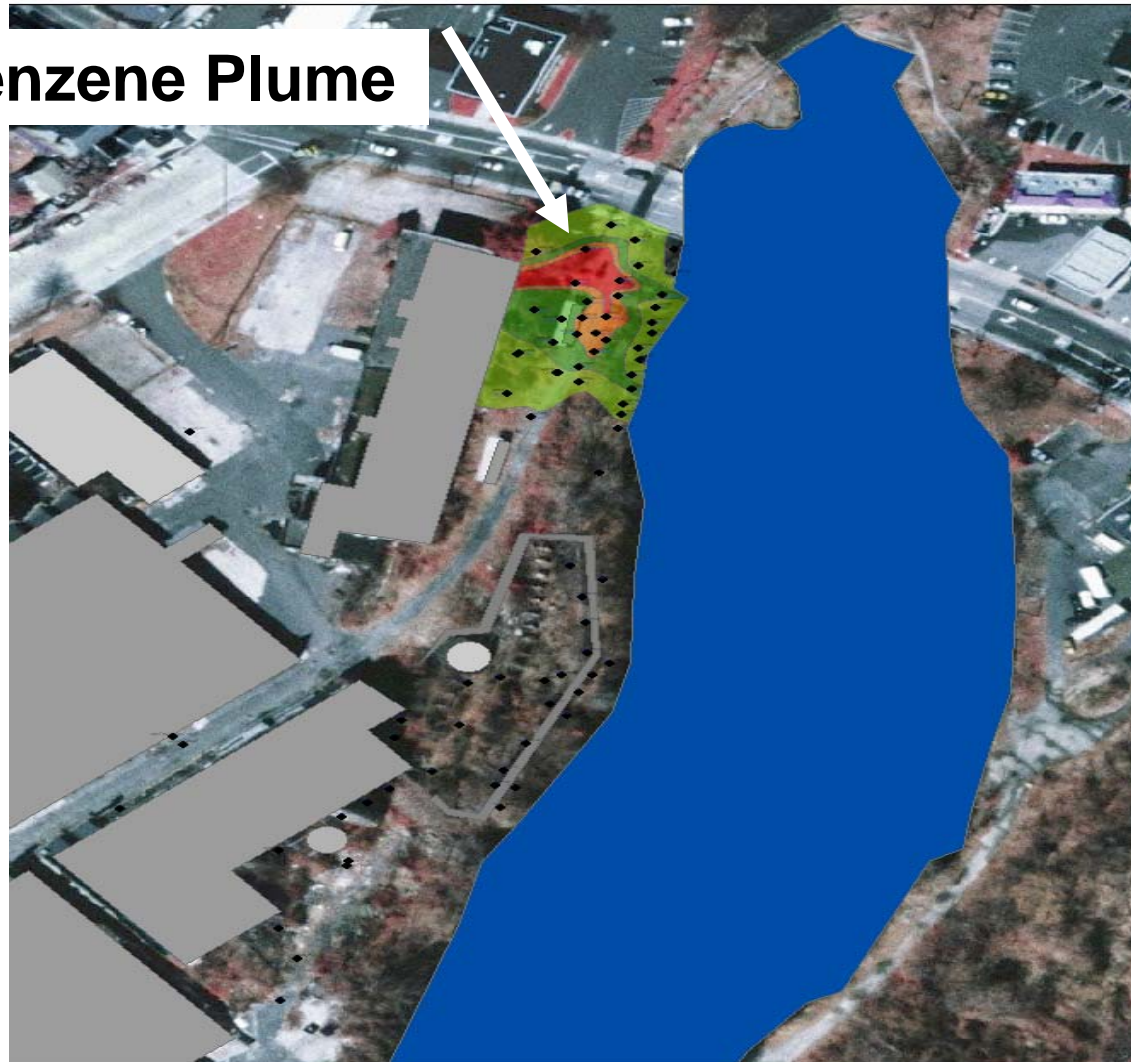# Stage 2 SI/RI Target Investigation Areas
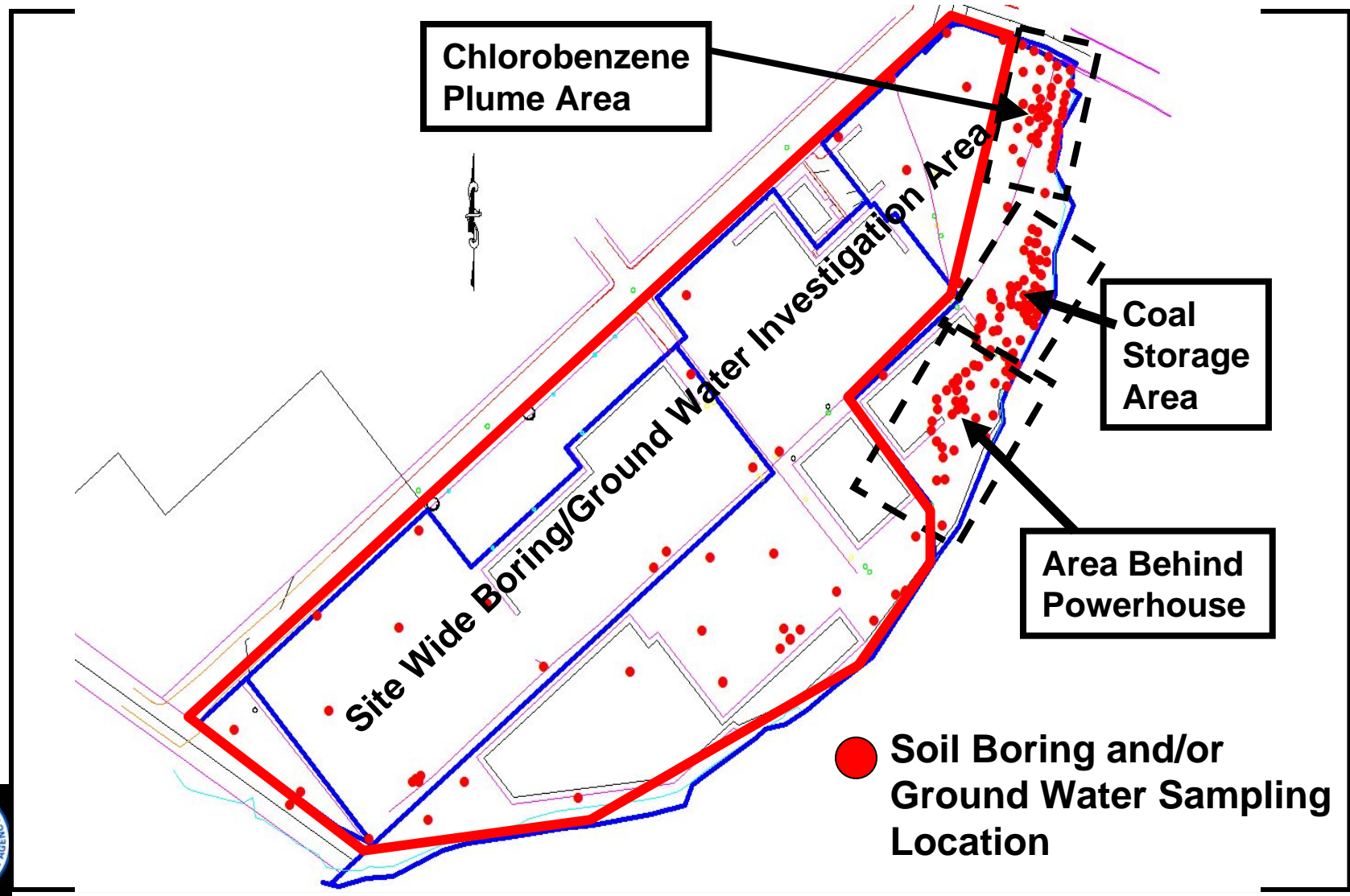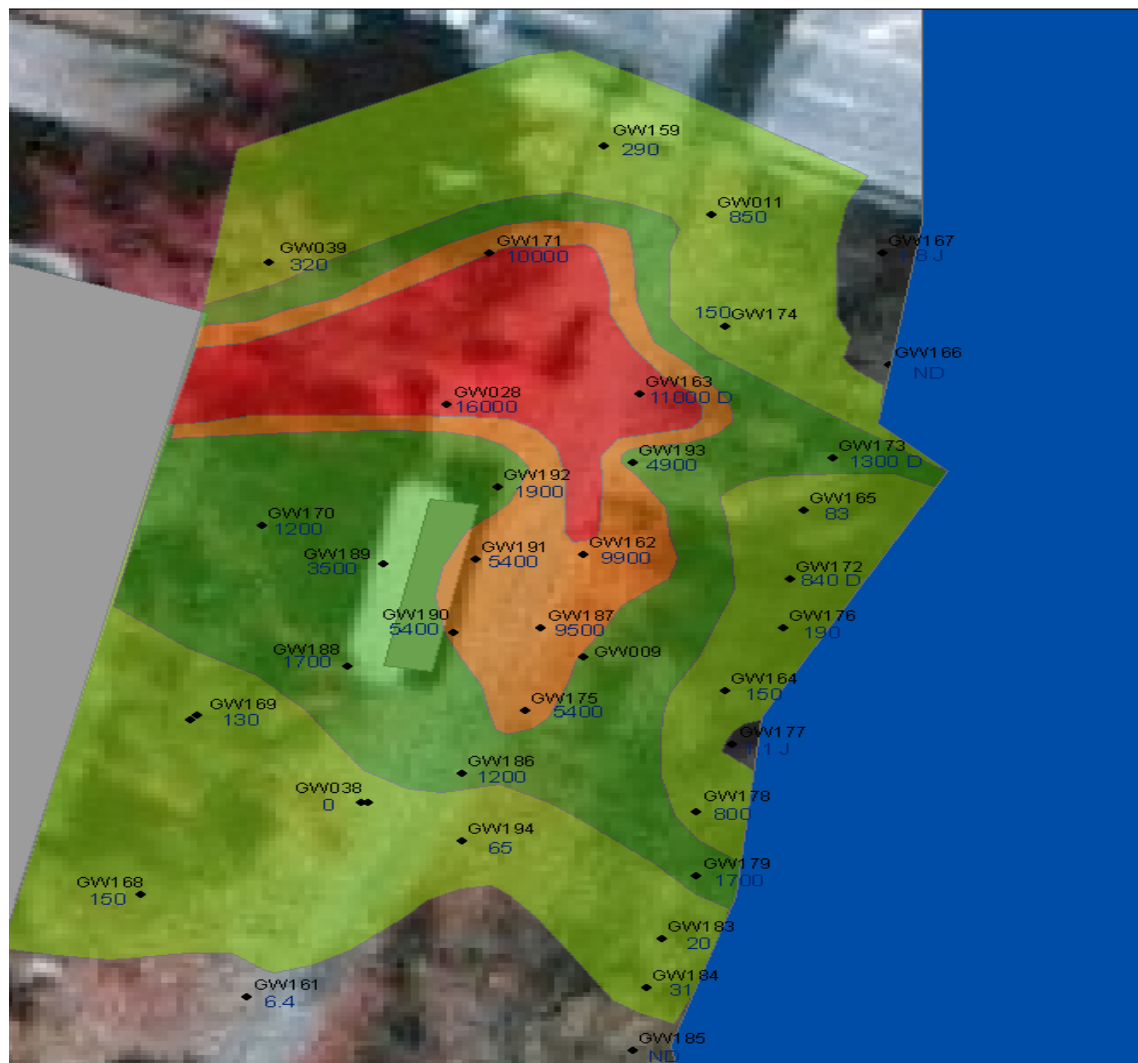
# Site Wide Soil Borings Eliminate Red Numbered Borings

# Location of Chlorobenzene Plume



Chlorobenzene Plume

# Chlorobenzene Plume Detail

# Milltown Redevelopment Project

◆ Lessons Learned

- » Scribe & Scriblets: allow for rapid processing of large amounts of collaborative field data

- » Web site very useful in allowing stakeholders to continuously update on findings

- » Chlorobenzene source investigation and plume delineation completed in a single mobilization

- » Field data useful in focusing where highest level data would be of the greatest value

# Site Information Management: Communication

# ConSoil 2008 – Milan

Carlos Pachon

U.S. EPA

Technology Innovation & Field Services Division

pachon.carlos@epa.gov

703-603-9904

# Value of Effective Data Communication
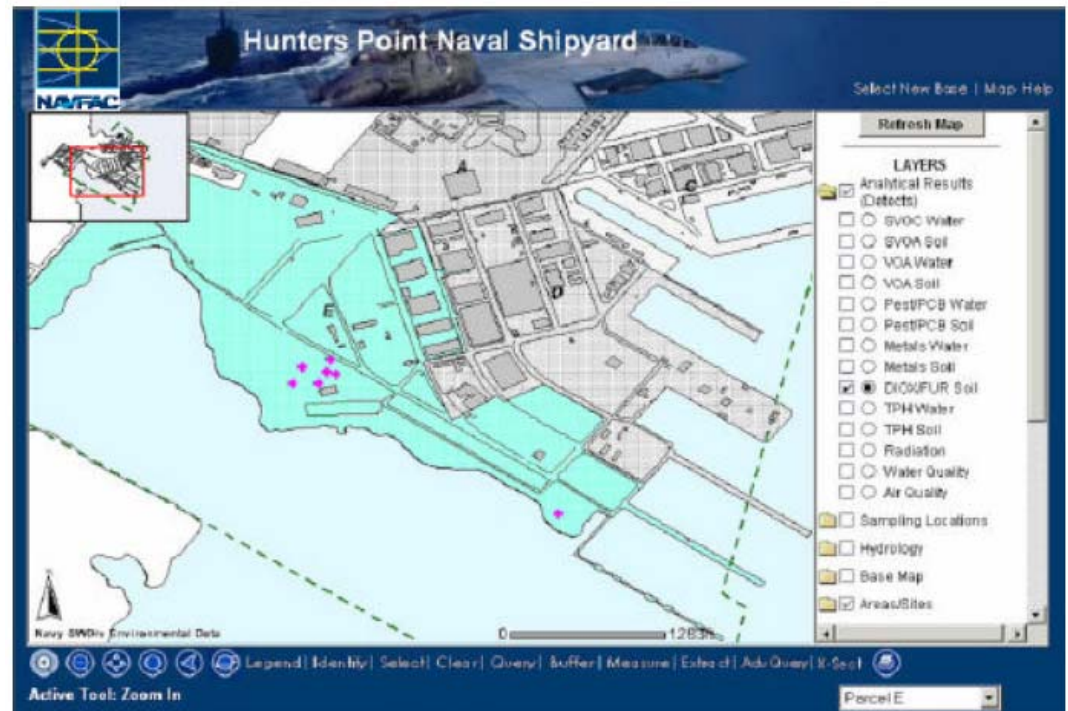
◆ Share results remotely with technical experts (better decisions)

◆ Engage all stakeholders more thoroughly

» Build client trust

» Facilitate regulator buy-in

◆ Facilitate team work and secure robust decisions

# Tools to Communicate Information

◆ Project/Team Websites

» Intra and internet

» Web portals

◆ Online Meetings

◆ Document Libraries

◆ Virtual Resources

# What is a Project/Team Web Site?

◆ A tool that can be used to _efficiently communicate_ with team members and _centrally manage information_ for a project

◆ Combines and expands upon commonly used resources such as:

» E-mails

» FTP sites

» Electronic calendars

» Task lists

◆ You can think of it as a "shared" drive on the Internet

# Why Use a Project/Team Web Site?

◆ Site teams or workgroups made up of diverse stakeholders:

  » Local governments

  » State and federal regulators

  » Contractors

  » Site owners and developers

◆ Projects involve many forms of information exchange:  reports, data, presentations, meetings, approvals, etc.

# Project/Team Web Sites

◆ A project/team website can provide:

  » Discussion – use in lieu of email threads

  » Library – store documents, reports, etc.

  » Calendar – schedule key meetings and events

  » Tasks – track actions and milestones

  » Members – keep contact information centrally located

◆ Options can be customized to meet the needs of a particular team

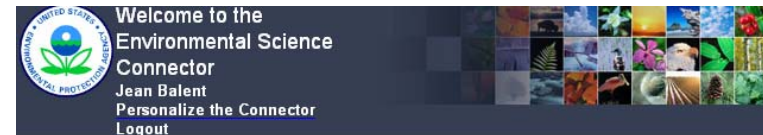◆ Site is created and managed by its members so there is virtually no expense to the project team

# Project/Team Web Site Tools

◆ EPA WebOSC



◆ EPA Environmental Science Connector (ESC) Workbench



◆ EPA Portal Collaboration Workspace



◆ Lotus QuickPlace

# Online Meetings

◆ Share materials live

◆ Share other applications

◆ Let other participants share materials

◆ Polling

◆ Chat with host, public, or groups of participants

◆ Record entire event for future download or replay

◆ Send automatic e-mails

◆ And more….

# Tools to Host Online Meetings

◆ EPA Portal Oracle Web Conference
  » http://hawkeye.epa.gov/imtapp/app/prelogin.uix
  » Public page, anyone can become a user
  » Can share anything live from reports, models, movies, etc.
  » Audio can be shared with a PC microphone
  » Used to walk people through presentations, view data live, review illustrations of site

◆ Sametime
  » https://epastx.rtp.epa.gov/epacenter.nsf
  » Username/password required
  » Can share anything live
  » Used to walk people through presentations, view data live, review illustrations of site

# Other Online Meeting Tools

◆ Go2Meeting

**GoToMeeting**® Online Meetings Made Easy™

◆ MeetingPlace

**CISCO**

Welcome to Cisco Unified MeetingPlace
* Meeting ID

◆ EPA Teleconference Web site

» http://www.teleconferencingcenter.com/

» Can share PowerPoint files online

# Other Communication Tools

◆ Instant Messaging

  » Sametime – built into Lotus Notes

  » Oracle Messenger – install
    and use from anywhere

◆ Document Library Sites

  » EPA Portal Content Services

  » SDMS – a Simple Document
    Management System for Lotus Notes
    (freeware)

◆ Virtual Tools

  » Mapping

  » Geospacial
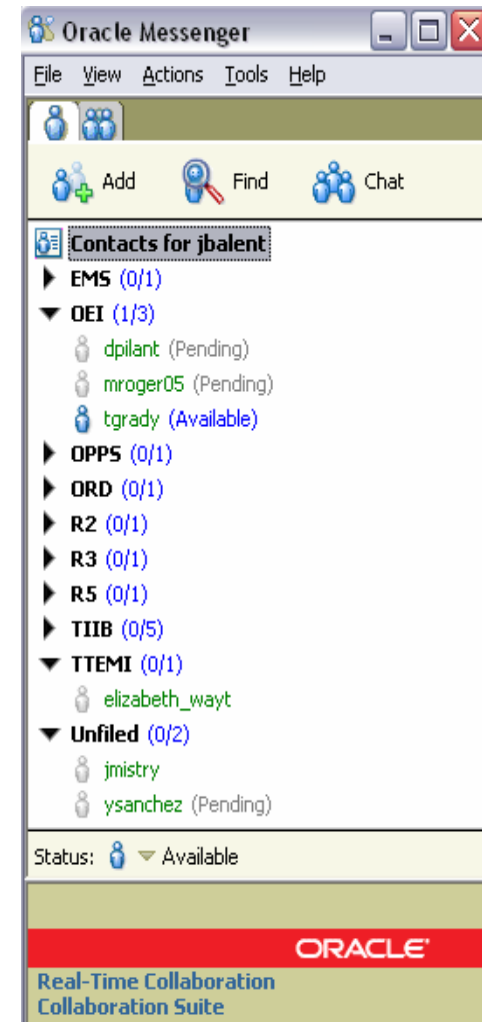
# EPA Mapping Resources

◆ http://www.epa.gov/enviro/

 » Runs Virtual Earth in background
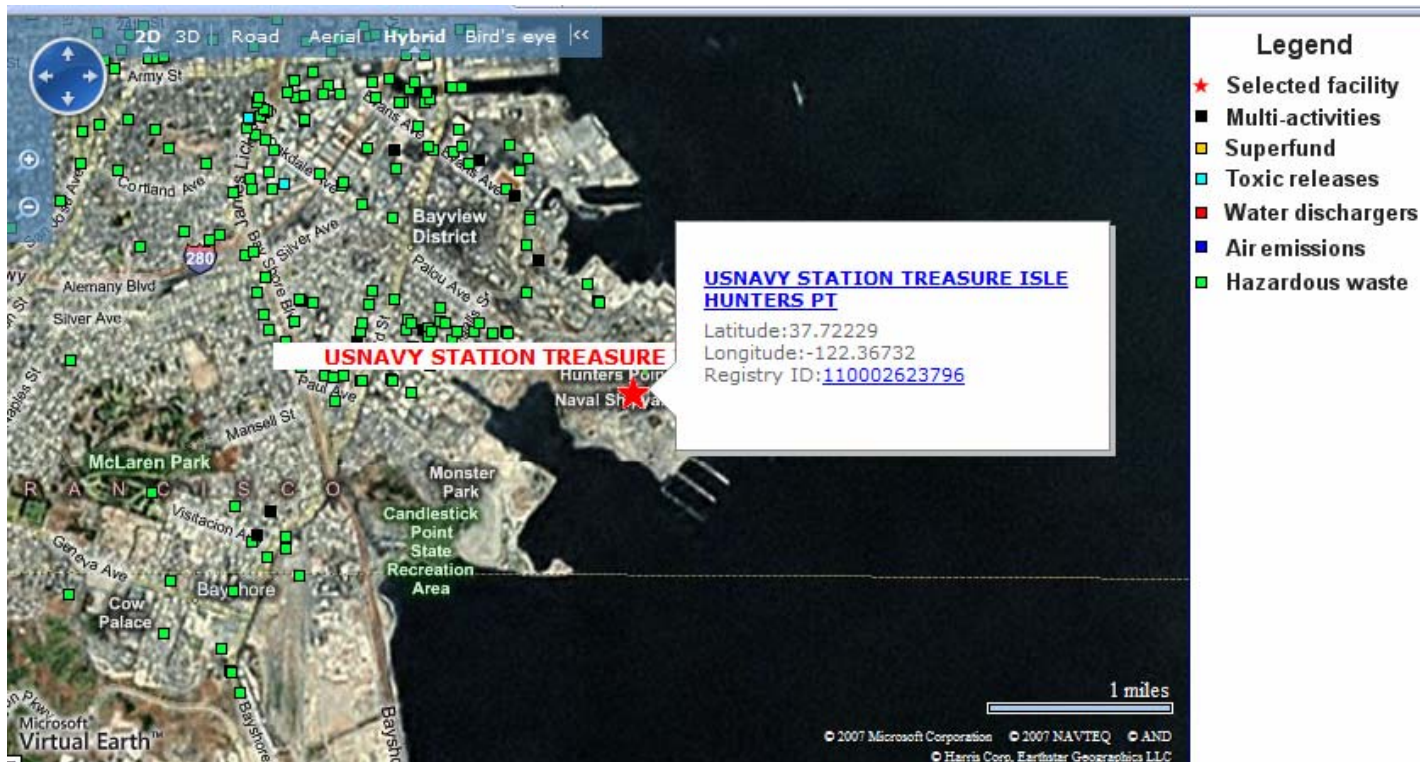
# Site Information Management: Make Decisions

# ConSoil 2008 – Milan

Tom Purucker

U.S. EPA Region 4

706-355-8123

Purucker.Tom@epa.gov

# Decision Support Tools

◆ Spatial Analysis and Decision Assistance (SADA) – Freeware

◆ FIeld EnvironmentaL Decision Support (FIELDS) – Freeware built on proprietary GIS (Arc)

◆ EVS and MVS – Proprietary
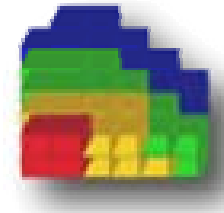
# SADA – A Decision Making Tool

- ◆ GIS
- ◆ Sample design
- ◆ Data management
- ◆ Statistics
- ◆ Data screening
- ◆ Geospatial interpolation
- ◆ Risk assessment (human, ecological)
- ◆ Uncertainty analysis
- ◆ Decision analysis
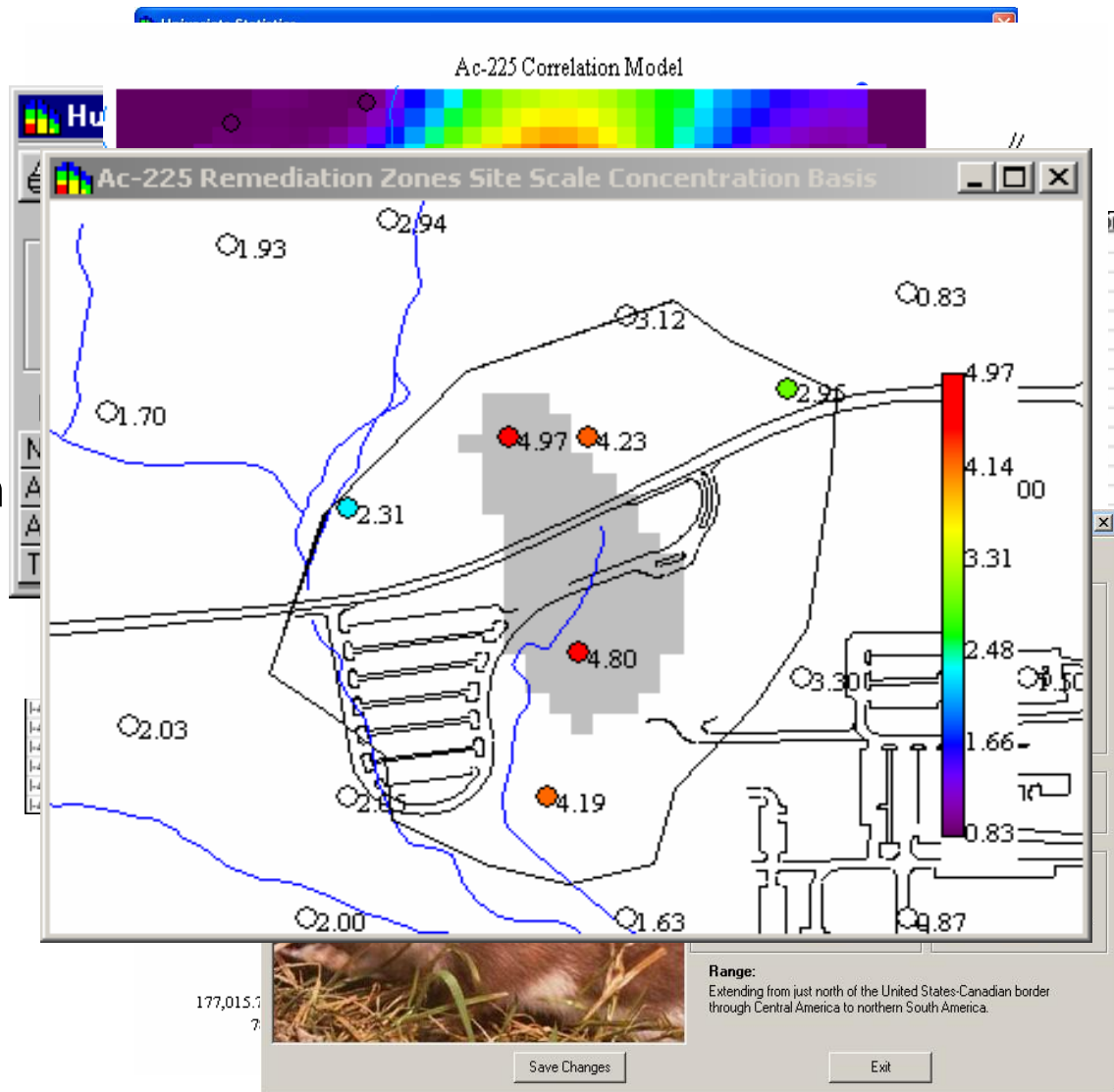
# SADA Capabilities and Decision Flow

Interpolate my data

**Interpolate my data**
Draw a variance map
Draw a probability map
Draw an area of concern map
Calculate cost vs cleanup
Draw a LISA Map
Develop sample design
Perform geostatistical simulation

## Steps

1. See the data
2. Set up the site
3. Set GIS overlays
4. Set grid specs
5. **Interpolation methods**
6. Choose helper data
7. Assess helper data
8. Correlation modeling
9. Search neighborhood
10. Show the results
11. Autodocumentation
12. Manage model results
13. Cross validation
14. Format picture
15. Export to file

<<Back    Help    Next >>

Show The Results

## Interpolators

Ordinary Cokriging          Help

**Modeling Options**
Geostatistics provides two options for estimating the value at any given point.

The mean is the kriging estimate for ordinary kriging and the E-type estimate for indicator kriging. This is the most popular choice for interpolation.

Percentile returns the values associated with the specified ccdf percentile.

⦿ Mean
○ Percentile          0.5
☑ Use this percentile for all data sets

**Type of Cokriging**
Intrinsic Coregionalization (Easy)

**Data Transform**
⦿ Unit transform (0 mean, variance 1)
○ No transform

**Intrinsic Model Variable for Correlation Modeling**
Primary

# XRF SADA Case Study

◆ Fort Lewis, Tacoma, WA - 2 former small arms ranges and a skeet range

◆ XRF data used to map soil volumes requiring treatment

◆ Volume estimates used to evaluate remedial alternatives

◆ Material > 1,000 ppm lead stabilized and reused
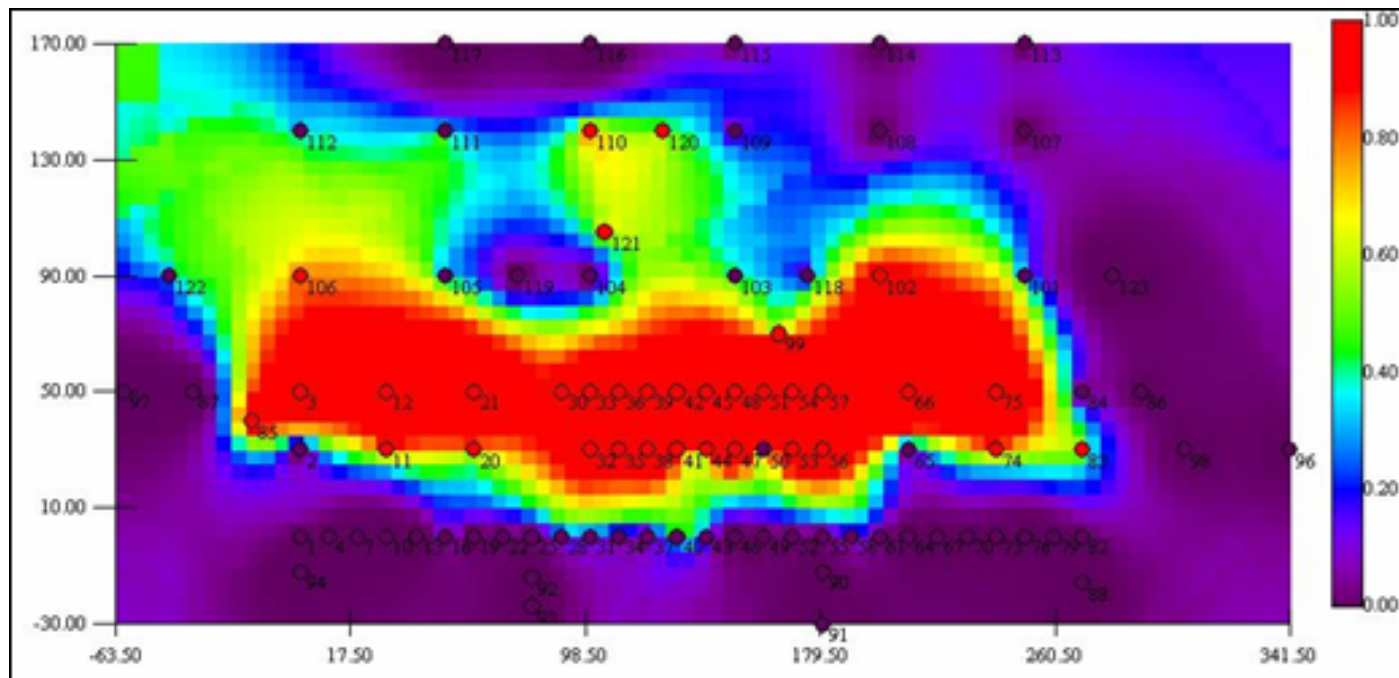
# FIELDS (FIeld EnvironmentaL Decision Support)

◆ **U.S. EPA Region 5 technical group**

◆ **Create software for contamination characterization**

◆ **Software includes**

  » FIELDS Tools for ArcGIS

  » F/S Plus

  » RAT (Rapid Assessment Tools)
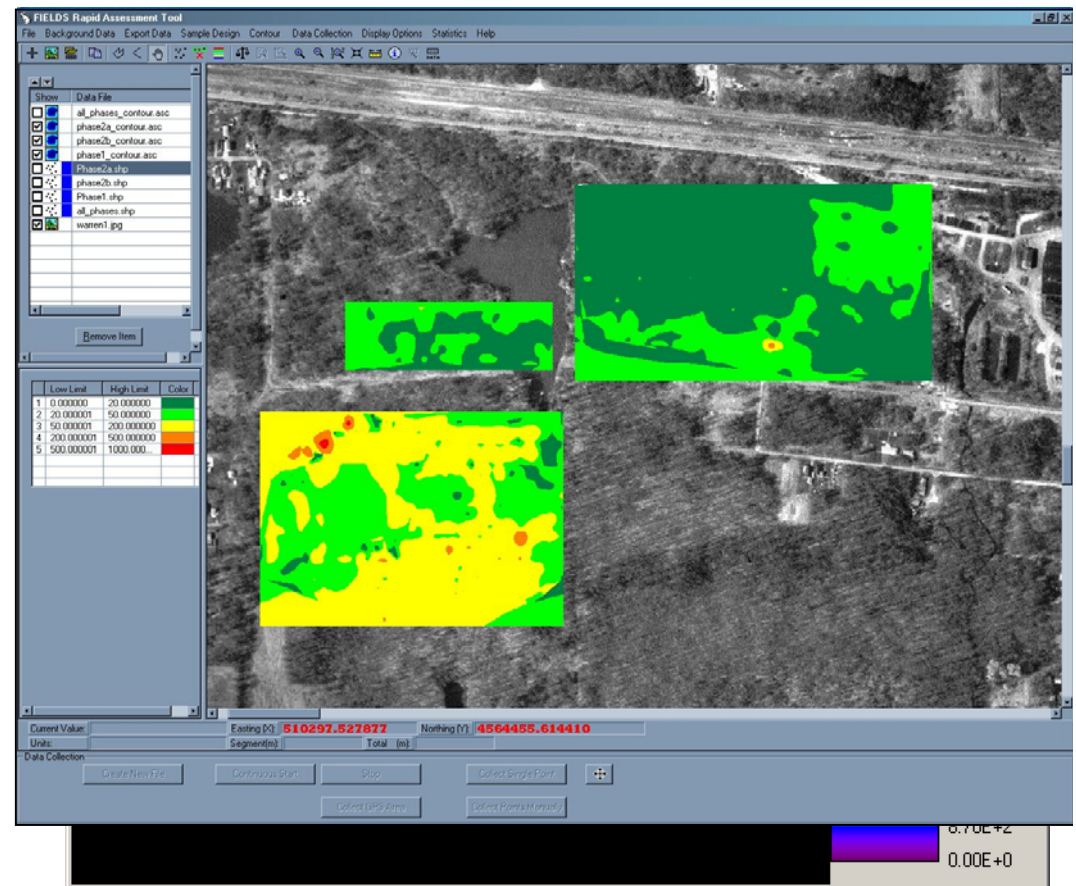
# FIELDS Tools

◆ The FIELDS Tools are a set of modules designed for contamination characterization. The modules include: sample design, database, modeling, and analysis tools (e.g., mass/volume estimation, remediation scenarios).

◆ The FIELDS Tools for ArcGIS can be downloaded at: http://epa.instepsoftware.com/FIELDS/
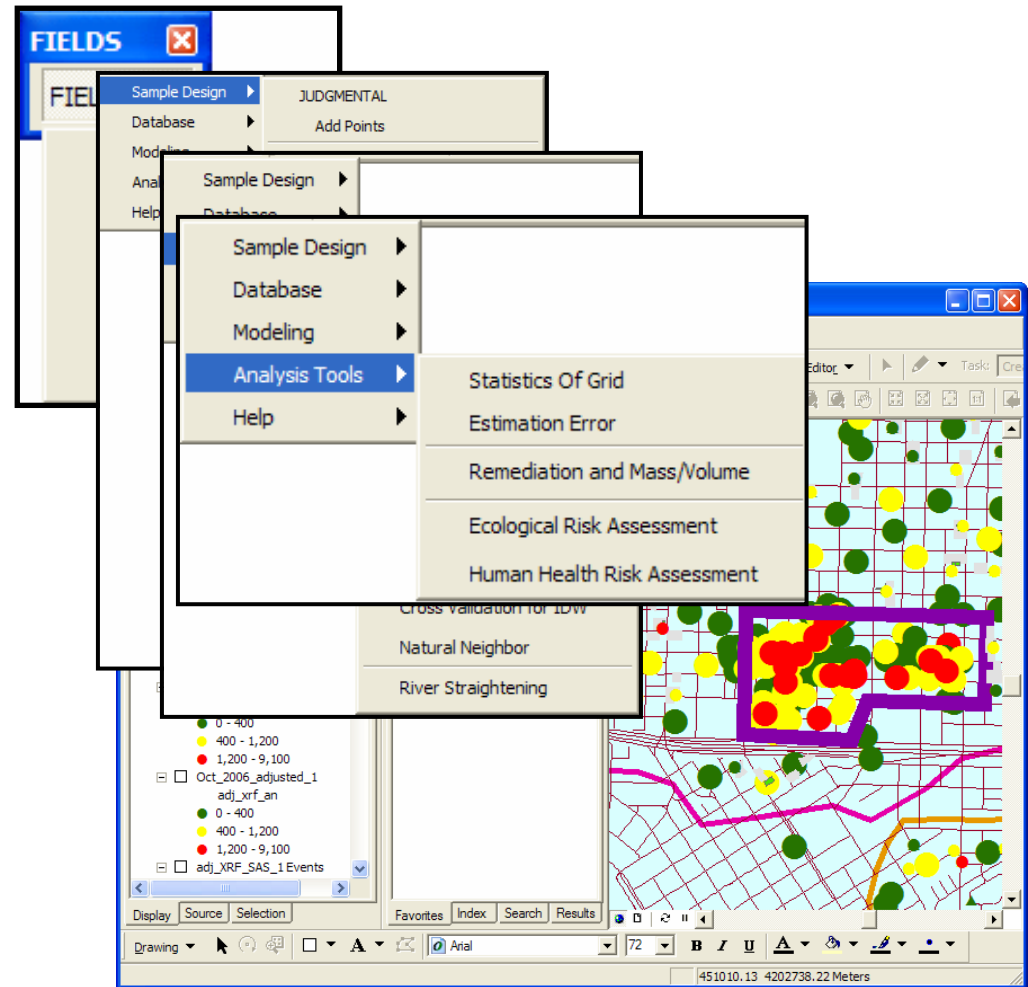
# F/S Plus

◆ F/S Plus is a stand-alone 2-D/3-D software developed in conjunction with SADA

◆ Allows for 2-D and 3-D modeling and analysis

◆ More information and to download F/S Plus: http://epa.instepsoftware.com/FIELDS/
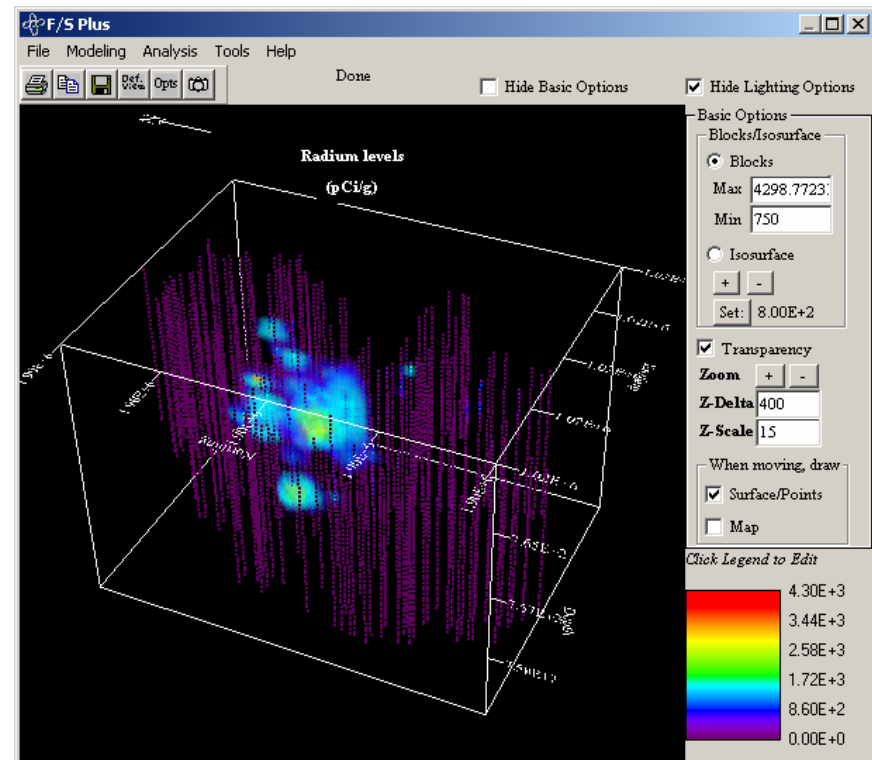
# Rapid Assessment Tools (RAT)

◆ RAT integrates real time GPS positions with data from external sensors (XRF, Rad units, Air monitors) to provide instantaneous snapshots of field conditions.

◆ Data are stored in GIS/Database compatible format eliminating need for conversion or manual data entry

◆ More information and to download RAT: http://epa.instepsoftware.com/RAT/
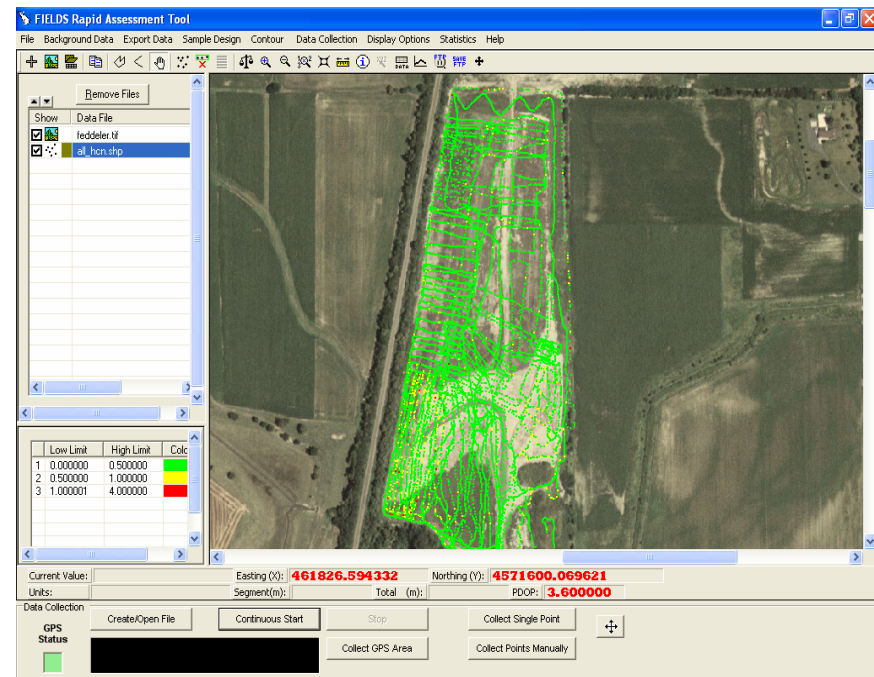
# FIELDS Software

Advantages:

◆ RAT allows in-field data collection and analysis

◆ RAT and F/S Plus are freeware

◆ FIELDS Tools and F/S Plus allow extensive contamination characterization

Limitations:

◆ FIELDS Tools require ArcGIS, proprietary

◆ Learning curve and time intensive to prepare graphics

# Case Study: Marino Bros. Scrapyard

◆ FIELDS was used to refine the preliminary CSM to guide the selection of an approach to site cleanup:

» query data

» develop ArcView shape files based on the query

» contour the data (isoconcentration maps)

» and perform mass and volume calculations

◆ The preliminary CSM was used to develop estimates of expected volumes of soil to be removed in order to clean up site for reuse

# FIELDS as a Visualization Tool



Aroclor 1254 in Soil
Site Specific Standard
0.0 to 2.0-foot
Interval

Marino Brothers
Scrapyard Site

Rochester, PA

Sample Point
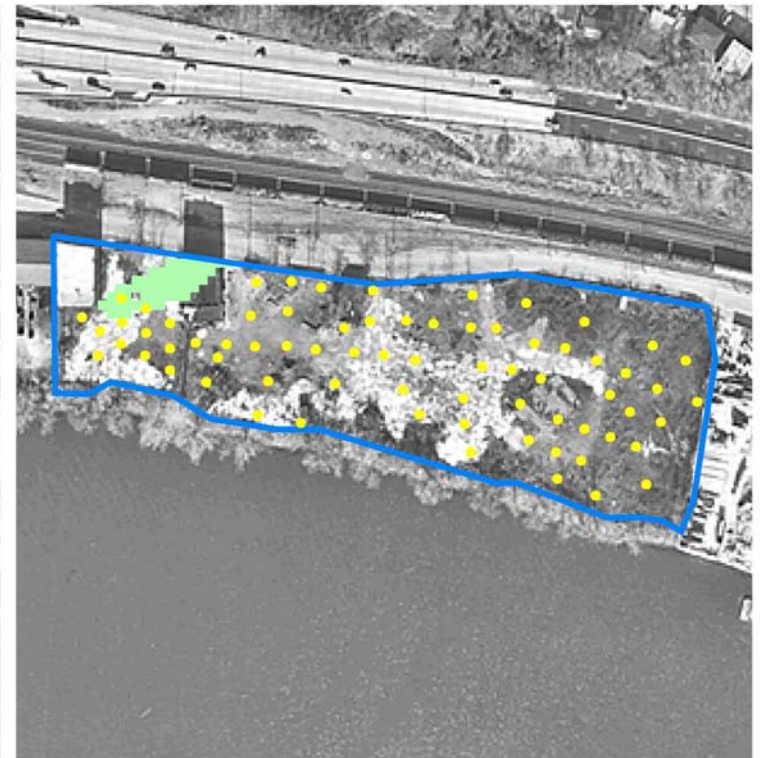0 to 2 feet bgs

Site boundary

Aroclor 1254 in mg/kg:

0 - 59*
59.1 - 590
590.1 - 5,900
5,900.1 - 33,009

Notes: Nondetects were assigned
the detection limit (practical
quantitation limit).

* Site-specific standard
for Aroclor 1254 in soil.

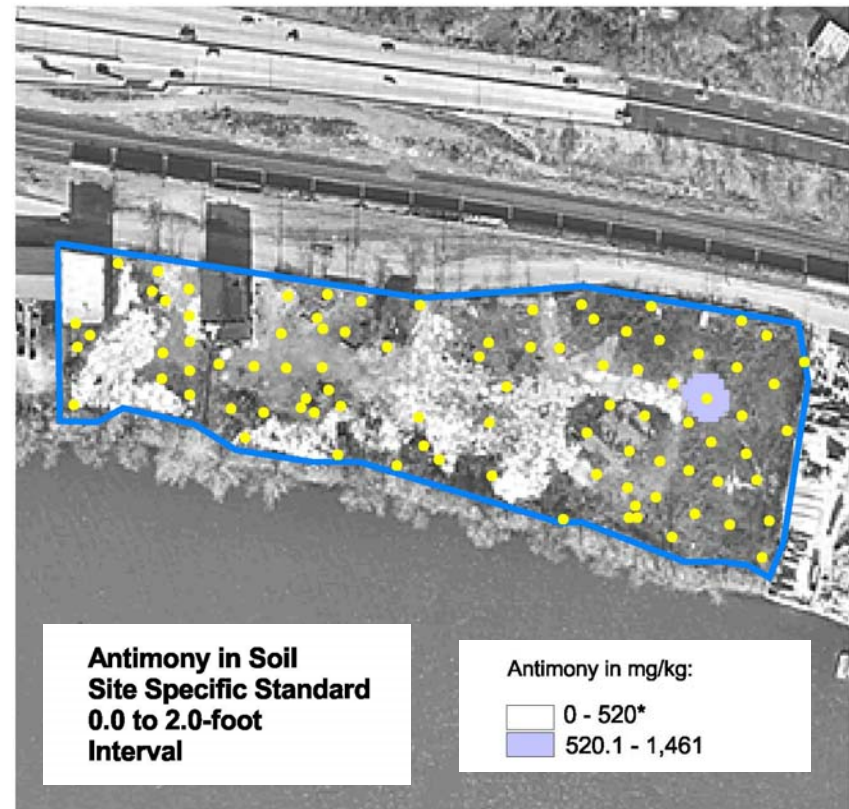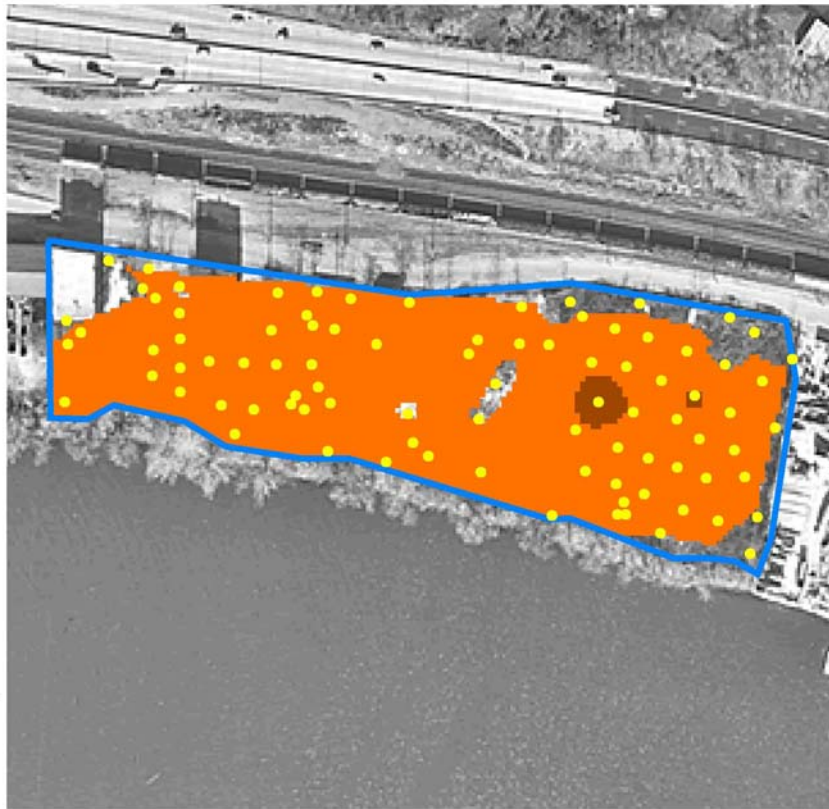0- to 2-foot interval          2- to 4-foot interval

# Comparing Extents of Contaminants



Isoconcentration maps highlight areas of overlap, allowing project team to focus on a smaller suite of analytes.

# Use of FIELDS at Marino Bros.: Outcomes

◆ Two types of estimates were needed to develop the SOW for site remediation: (1) volume of soil to be excavated, and (2) volumes of soil to be shipped to Class I and Class II landfills

◆ SADA was used in combination with FIELDS to meet project objectives

# EVS Software

Advantages:

◆ Proprietary software capable of 3-D visualizations and multiple views

◆ Time series depictions to identify trends

◆ Model compatible solids

Limitations:

◆ Cost

◆ Learning curve

# Lone Butte Industrial Park
# Case Study

◆ Large industrial park up gradient of a casino, golf course, and agricultural users in Phoenix, New Mexico

◆ Key Study Question?

» What will be the impact of a chlorinated solvent plume on increased water use down gradient of the site?

» Where are the sources of contamination?

# Site Geology with Logs



Site_Geology_w_Logs
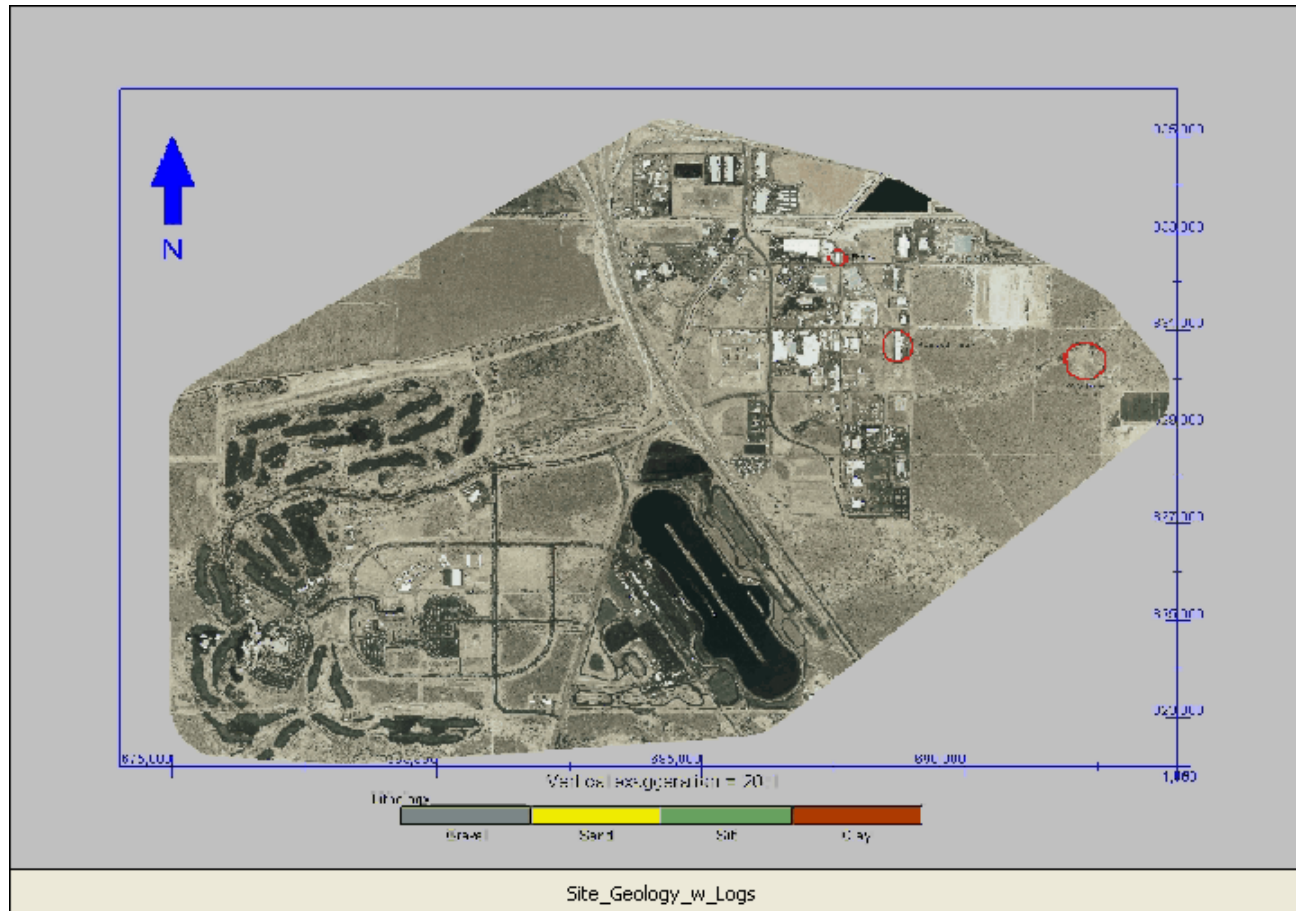
# TCE Oblique View
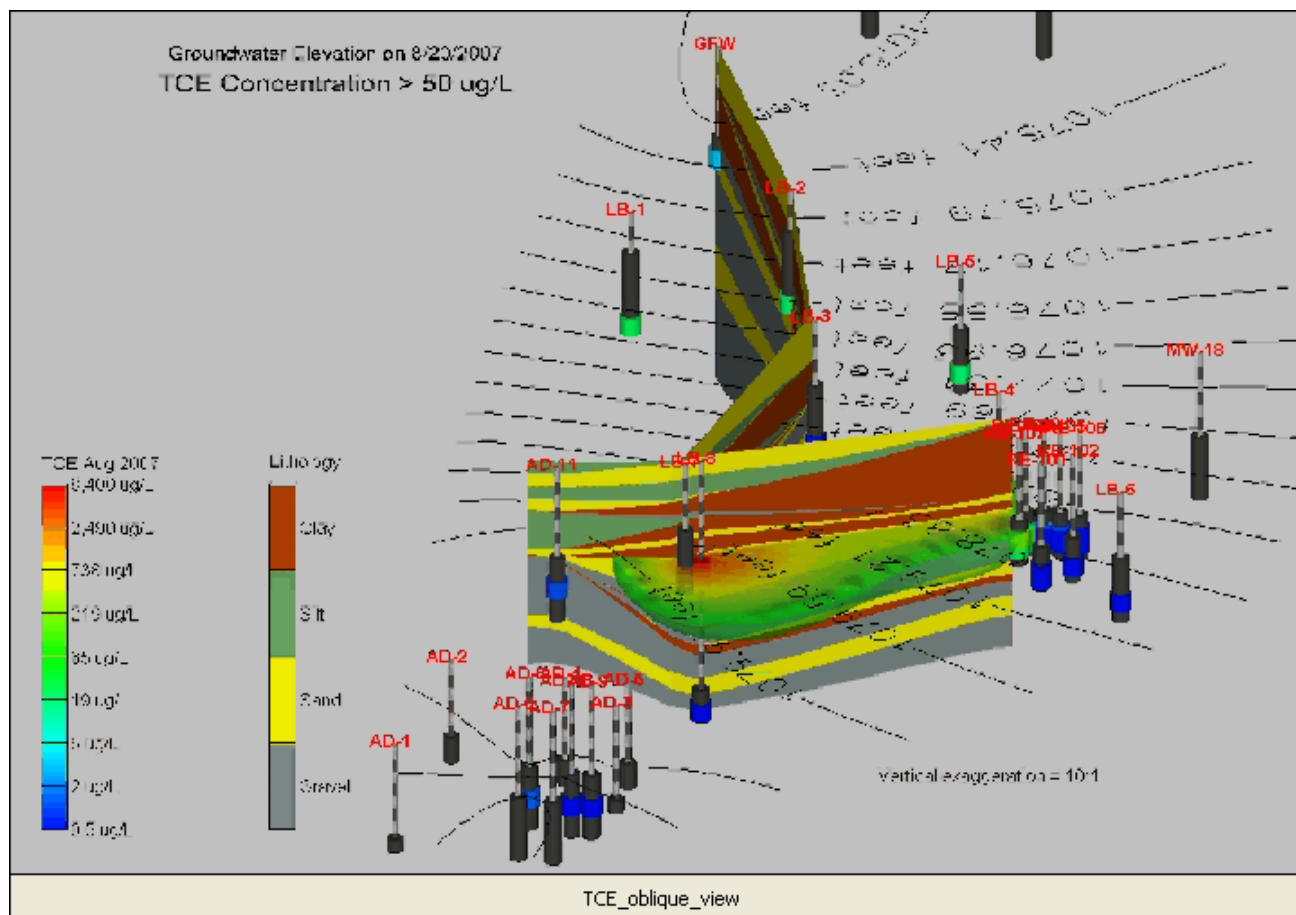
# Lone Butte Industrial Park Summary

◆ Data visualizations were used to identify potential source areas

◆ Visualizations showed how geology will control migration pathways

◆ Critical data gaps in results were identified

◆ Future sampling events were optimized

# Summary

◆ Triad projects benefit from the use of nimble real-time data collection, processing, and communication tools to support decision making

◆ Real-time decision making can dramatically reduce overall project costs and improve the certainty with which decision are made

◆ More data evaluation and storage tools are becoming available, but project teams need to plan ahead to be successful

# Thank You!
## Questions and comments are welcomed!

For more information

Carlos Pachon
U.S. EPA
Technology Innovation & Field Services Division
pachon.carlos@epa.gov
703-603-9904