



**Pacific Northwest**  
NATIONAL LABORATORY

*Proudly Operated by Battelle Since 1965*

# Incremental Sampling Methods: Current VSP Applications for Representative Soil Sampling

JOHN HATHAWAY, MARIANNE WALSH, MICHAEL WALSH

Pacific Northwest National Laboratory

2015 Environmental Data Quality Workshop (Portland, OR)

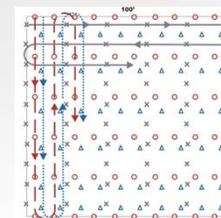
## Full Disclosure

- ▶ I am a statistician
- ▶ Graduate thesis was on composite sampling
- ▶ I had many concerns about multi-increment sampling (MIS) when I was introduced to it in 2006
- ▶ Participated on the Interstate Technology & Regulatory Council (ITRC) committee that drafted guidance on the Incremental Sampling Methodology (ISM)
- ▶ Currently funded to integrate statistical sample design concepts into MIS/ISM in collaboration with the original developers (Cold Regions Research and Engineering Laboratory)
- ▶ I have no preference regarding the naming of a formalized compositing process



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



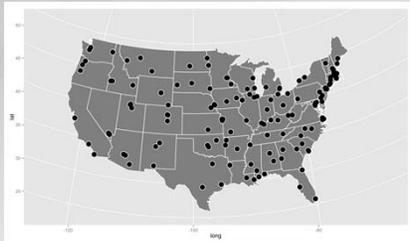
# Outline



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965

- 1) Simple example of representativeness



- 2) Review of Multi-Increment Sampling (MIS), Incremental Sampling Method (ISM), and composite sampling

- 3) Background of Visual Sample Plan and MIS/ISM



- 4) Field demonstration of statistically driven ISM designs



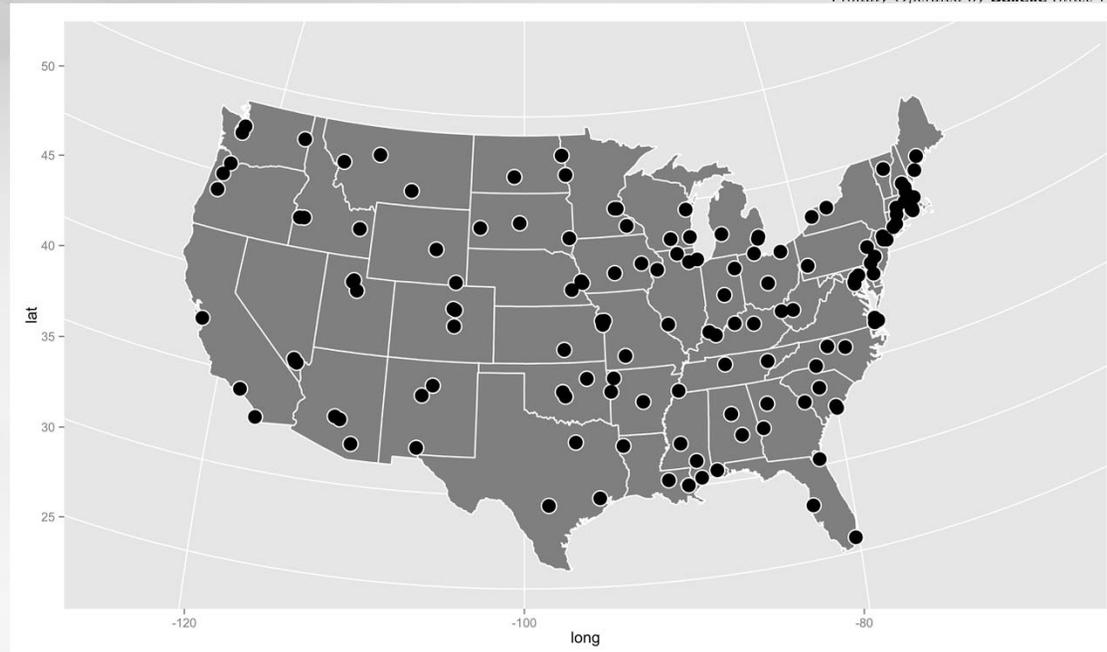
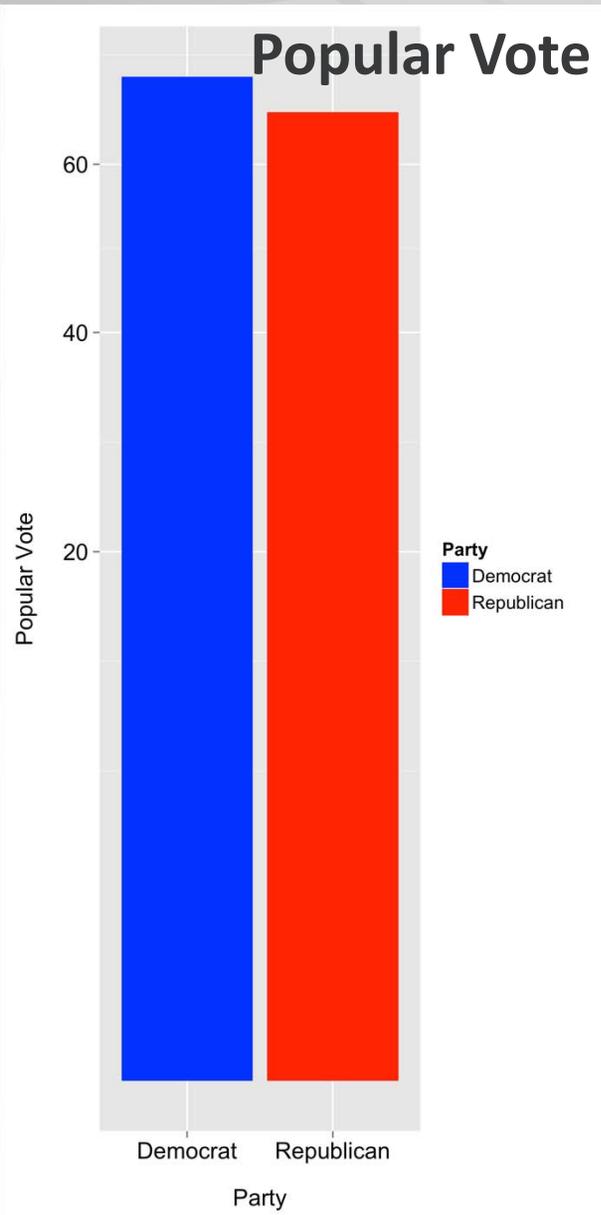
**R**epresentative  
e**pr**oducible  
a**nd**omized

# Decision Units & Presidential Elections



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle**. Since 1965



- ▶ Random by state sample: Polled 144 people
- ▶ Democratic candidate has 52% of the vote
- ▶ Thus, the democratic candidate is the most likely to win
- ▶ 5% standard error

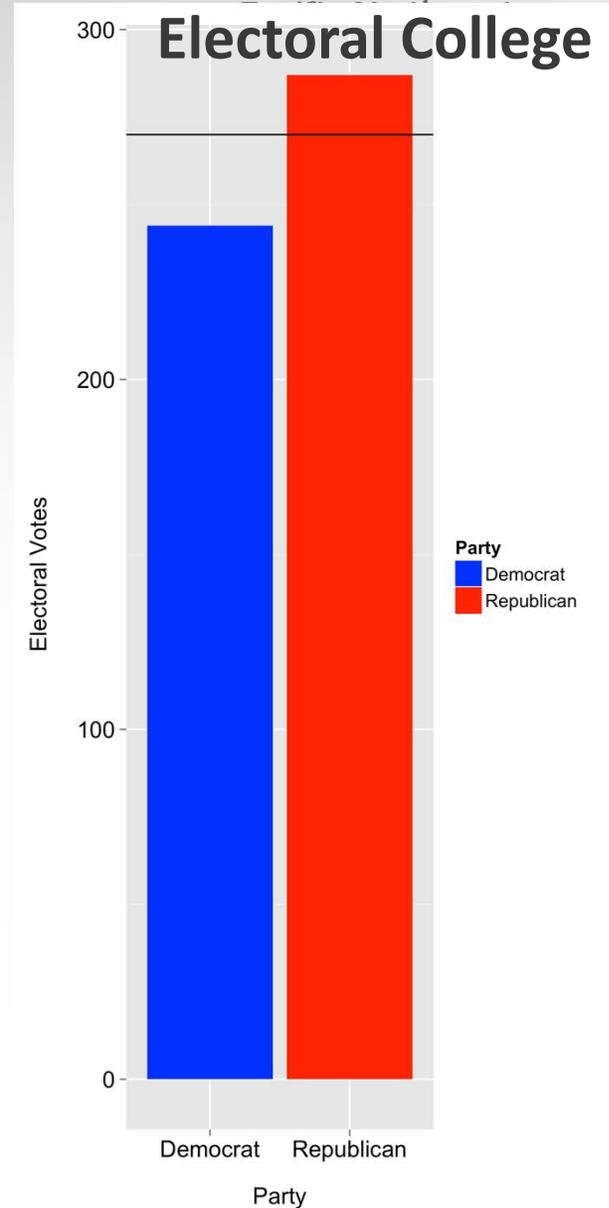
# Decision Units & Presidential Elections



But elections are not decided by the popular vote.

Now what do we do?

We have sampled 3 people for every state. We can estimate which way each state will go and tally up the electoral votes.





# What is Incremental Sampling?



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965

## ▶ Environmental Sampling Goal:

Obtain representative data for decision making on a study area

- ▶ Incremental sampling is the process of establishing pertinent decision units and then collecting multiple increment samples that are

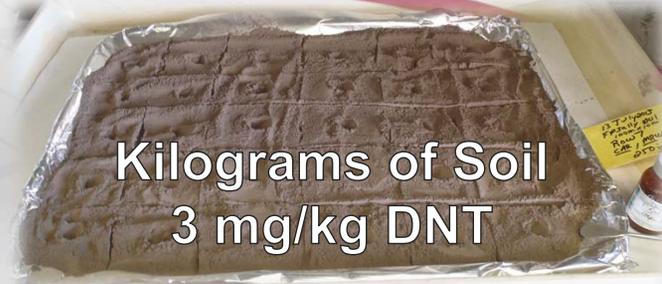
**R**epresentative  
e**pr**oducible  
a**nd**omized

- ▶ ISM forces the conversation about “representativeness”
  - One sample from 2 acres most likely is not representative of the 2 acres
  - Too much grouping/compositing may underrepresent/dilute elevated regions within a wide area



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



Kilograms of Soil  
3 mg/kg DNT

Tons of Soil  
3 mg/kg DNT



**Representativeness:  
Field Sample  
Collection to  
Lab Analysis**



10 g of soil  
3 mg/kg DNT

100  $\mu$ L of solvent analyzed



Solvent  
Extraction  
0.03 mg  
DNT

# DOE Funded Design Development

## ► Methods development research and integration into VSP (2008-2012)

Look Inside Get Access

Environmental and Ecological Statistics  
March 2013, Volume 20, Issue 1, pp 69-90

Date: 19 Jun 2012

### Cross-combined composite sampling designs for identification of elevated regions

John Hathaway, Stephen Walsh, Landon Sego, Brent Pulsipher



Confidence Interval on True Mean - Multiple Increment Sampling

Confidence Interval | Sample Placement | Costs | Analytes |

For Help, highlight an item and press F1

Analyte:

I want to be 95.00 % confident that the estimated mean is within 5 units  
above or below the true mean. (Two-sided confidence interval)

The estimated standard deviation due to increment-to-increment variability is 25 units.

The estimated standard deviation due to the analytical method is 1 units.

I expect to conduct 1 analyses per sample. Fixed Using best values for lowest cost (see Costs tab)

Each MI sample should be composed of 71 increments Fixed

Minimum Number of MI Samples for Analyte 1: 4

Minimum Number of Multiple Increment (MI) Samples: 4

If 1 analysis is performed on each multiple increment sample composed of 71 increments, then 4 of these multiple increment samples are required to achieve 95% confidence that the true mean lies within the confidence interval half-width of the estimated mean.

A total of 284 increment samples will be placed on the map.

OK Cancel Apply Help

Detecting Elevated Contamination Using Multiple Increment Sampling (beta)

MI Sampling | Increment Placement and Ordering | Data Entry and Analysis | Analytes |

Null Hypothesis: Elevated constituent concentration does not exist in areas represented by an increment.  
Alternative Hypothesis: An elevated constituent concentration exists in an area represented by an increment.

These design parameters apply to Analyte 1

Specify Parameter to Calculate:

I want to calculate:

the minimum size elevated area that can be detected and the maximum permissible number of increments in each row and column.  
 the probability of detecting an elevated concentration given a specified number of increment samples.

Specify Design Parameters:

I will have at least a 91.2315 % probability of detecting (power) an area with contamination levels greater than 6 standard deviations greater than background

If no elevated areas exist, then I only want a 5.0 % chance (alpha) of identifying an area as elevated.

A maximum\* of 7 rows and columns can be used to meet the prescribed detection objectives.

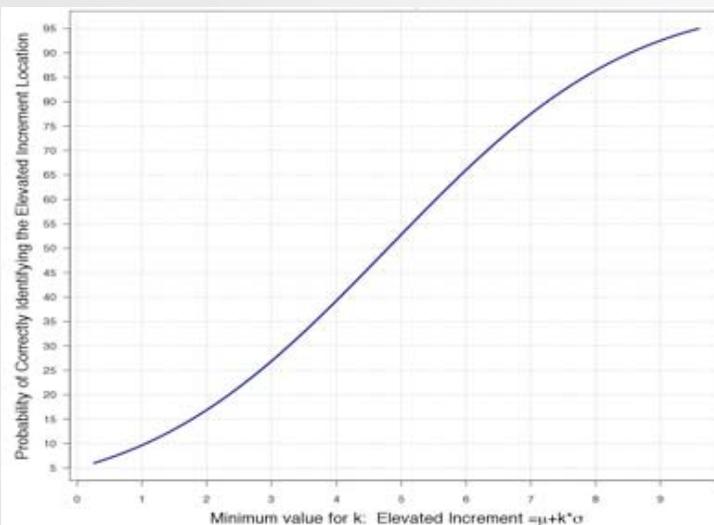
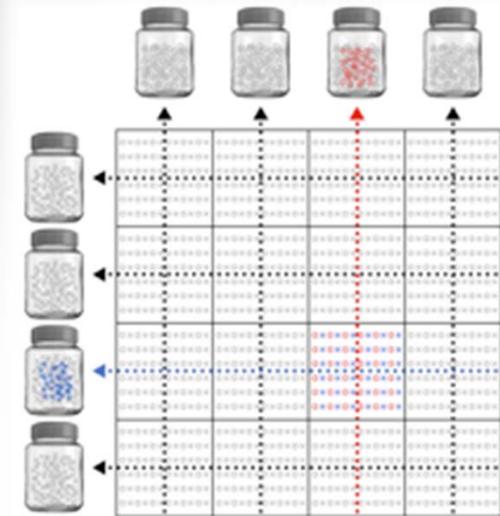
Thus, the smallest detectable elevated area will be 255.764 square meters

If each increment is assumed to be representative of the surrounding 9.57 x 26.73 meter region of 255.764 meters<sup>2</sup>, for a grid with 7 rows and columns, there will be 49 total increments placed to be combined into 14 total MI samples and the probability of detecting an elevated area that is more than 6 standard deviations above the background mean will be at least 91.23%.

OK Cancel Apply Help

# Technical Description: Aligned ISM Sampling Designs

- The Aligned MI design -- The ability to collect a reduced number of ISM samples over large areas while still preserving information about the spatial distribution of the contaminant
- PNNL developed the theoretical model to quantify the probability that a single elevated grid cell could be identified using a pre-defined “row” and “column” ISM sampling configuration



$$k = 0.5\sqrt{r + c + 2} (Z_{1-\alpha} - Z_{1-\pi})$$

r number of rows

c number of columns

k # of standard deviations above background.

$(Z_{1-\alpha} - Z_{1-\pi})$  Type I & II Error

# Technical Description: ISM Designs for Decisions on Mean



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965

- Provide a **site specific DQO-based** answer to **how many increments** should be combined in each ISM sample and **how many ISM samples** should be taken within a wide area decision unit (DU)
- Three designs for mean based DQOs
  - Compare to a threshold
  - Construct a confidence interval
  - Compare to background
- Methods adapted from published paper (Rhoff, 1996)

**True Average vs. Fixed Threshold**

Average vs. Fixed Threshold | Sample Placement | Costs | Analytes

I  can assume the data will be normally distributed. For Help, highlight an item and press F1

I want to use  multiple increment sampling.

These design parameters apply to

**Specify Null Hypothesis:**  
I want to assume the site is  until proven otherwise.  
(Assume the true mean  $\geq$  action level.)

**Specify False Rejection Rate (alpha) and Action Level:**  
I want at least  % confidence that I will conclude the site is unacceptable (dirty) if the true mean is at or above the action level of  units.

**Specify Width of Gray Region (delta) and False Acceptance Rate (beta):**  
If the true mean is  units below the action level (that is, 0.4 units) then I want no more than a  % chance of incorrectly accepting the null hypothesis that the site is unacceptable (true mean  $\geq$  action level).

**Specify Multiple Increment Sampling Options:**  
The estimated standard deviation between increments is  units.  
The estimated standard deviation between analytical subsamples is  units.  
The number of increments in each Multiple Increment (MI) sample  and the number of analytical subsamples taken from each MI sample

To optimize for  each MI sample will consist of  increments and have  analytical subsamples taken.

Minimum Number of MI Samples for Analyte 1: 3

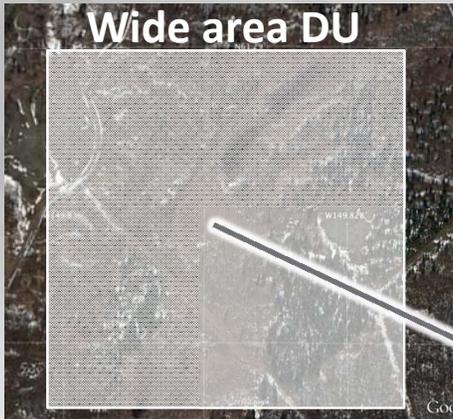
Minimum Number of Samples in Survey Unit: 3

# Technical Description: Wide Area ISM Designs



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



Very large DUs  
restrict usability  
of current full  
DU ISM  
samples

Data Quality  
Objectives  
(DQO)

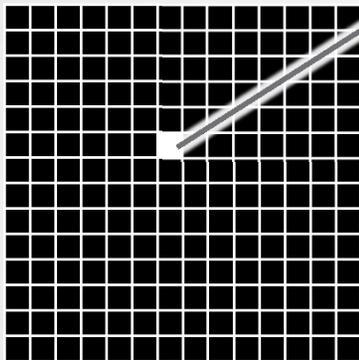


Statistical sampling  
designs can leverage  
small sampling units  
to make inference to  
the wide area DUs

## DU mean decisions:

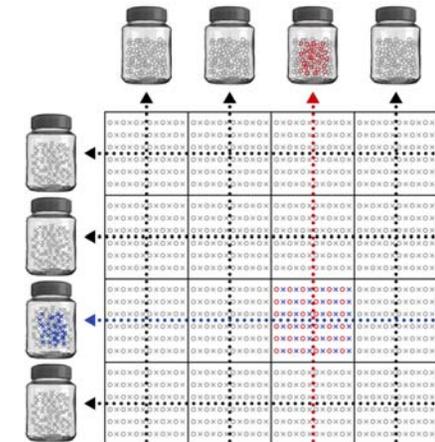
VSP tools use traditional  
statistical sample size  
equations adapted to ISM to  
define how many ISM  
should be collected and the  
number of increments to  
include in each.

~Hundreds of 1/4 acre  
sampling units



Collecting  
hundreds of  
ISM  
samples  
is cost  
prohibitive

## Dispersion decisions:



# ESTCP Funded Demonstration Project

- ▶ Demonstrate the integration of statistically based ISM designs within VSP that incorporate ISM/MIS methods previously demonstrated under SERDP/ESTCP for wide area applications (2013-2016)



The screenshot shows the SERDP/ESTCP website interface. At the top, there are logos for SERDP (DOD, EPA, DOE) and ESTCP, with the tagline 'Meeting DoD's Environmental Challenges'. A search bar and social media icons are also present. The main navigation menu includes Home, About SERDP and ESTCP, Program Areas, News and Events, Featured Initiatives, Tools and Training, Funding Opportunities, and Investigator Resources. The 'Program Areas' sidebar lists various categories, with 'Identifying and Evaluating Sources' highlighted. The main content area displays the title 'Optimal Incremental Sampling Methods: Tools for Mean Estimation and Spatial Delineation' for project ER-201329, along with an 'Objective' section and a 'Points of Contact' sidebar.

**Program Areas**

- Energy and Water
- Environmental Restoration
- Contaminated Sediments
- Contaminants on Ranges
- Identifying and Evaluating Sources**
- Characterizing Fate and Transport
- Protecting Groundwater Resources
- Assessing Potential Ecological Impacts
- Contaminated Groundwater
- Wastewater and Drinking Water
- Risk Assessment
- Munitions Response
- Resource Conservation and Climate Change

[Home](#) > [Program Areas](#) > [Environmental Restoration](#) > [Contaminants on Ranges](#) > [Identifying and Evaluating...](#) > ER-201329 Project Overview

## Optimal Incremental Sampling Methods: Tools for Mean Estimation and Spatial Delineation

ER-201329

[Objective](#) | [Technology](#) | [Benefits](#)

### Objective

Cost-effective range characterization is a challenge for range management and remediation, particularly when characterizing the spatial distribution of munitions constituents (MC) over large areas. Incremental sampling methods (ISM) are often required within the Department of Defense (DoD) for characterizing MCs on ranges. Recent projects funded by SERDP and ESTCP have improved the sample handling and laboratory process necessary for ISM. The objective of this project is to address a current need for the application of statistical tools in a systematic planning process, such as the U.S. Environmental Protection Agency's (EPA) Data Quality Objectives (DQO) process, to optimize incremental sampling approaches over large areas. The statistically defensible sampling strategy to be demonstrated will allow DoD to control the risk of incorrect statistical sampling decisions that inherently exist and have a demonstrated strategy for use on military ranges.

[Back to Top](#)

**Points of Contact**

**Principal Investigator**  
Mr. John Hathaway  
Pacific Northwest National Laboratory  
Phone: 509-372-4970  
Fax: 509-375-2604  
[john.hathaway@pnnl.gov](mailto:john.hathaway@pnnl.gov)

**Program Manager**  
Environmental Restoration  
SERDP and ESTCP  
[er@serdp-estcp.org](mailto:er@serdp-estcp.org)

**Document Types**

**Project Overview** - Brief project summary with links to related documents and points of contact.

**Final Report** - Comprehensive report for every completed SERDP and ESTCP project that contains all technical results.

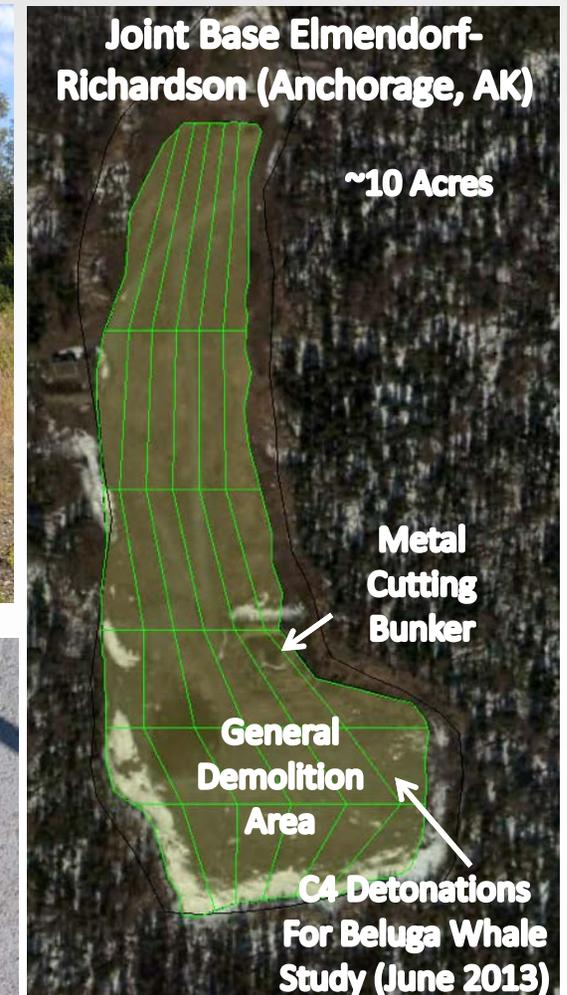
# Technical Approach

## 2014 – Statistical Sampling Validation within Wide Area DUs



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



# Technical Approach

## 2013 – Statistical Sampling Validation within Small Area DUs

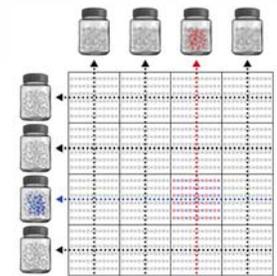
- Compare mean estimates from VSP designs to traditional ISM designs
- Validate robustness of aligned ISM assumptions

### Four Design Scenarios

#### VSP Mean Estimation Design

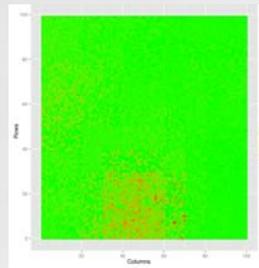


**>400 laboratory analyses during phase 2**

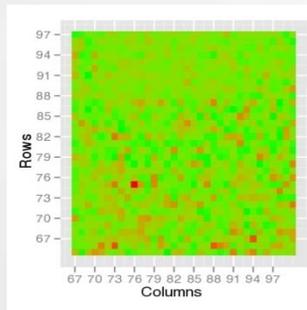


**Aligned ISM  
10 x 10 grid**

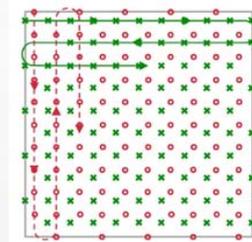
Two DU Scenarios



DU 1



DU 2



**Traditional ISM  
100 increments**

# Alaskan Training Ranges

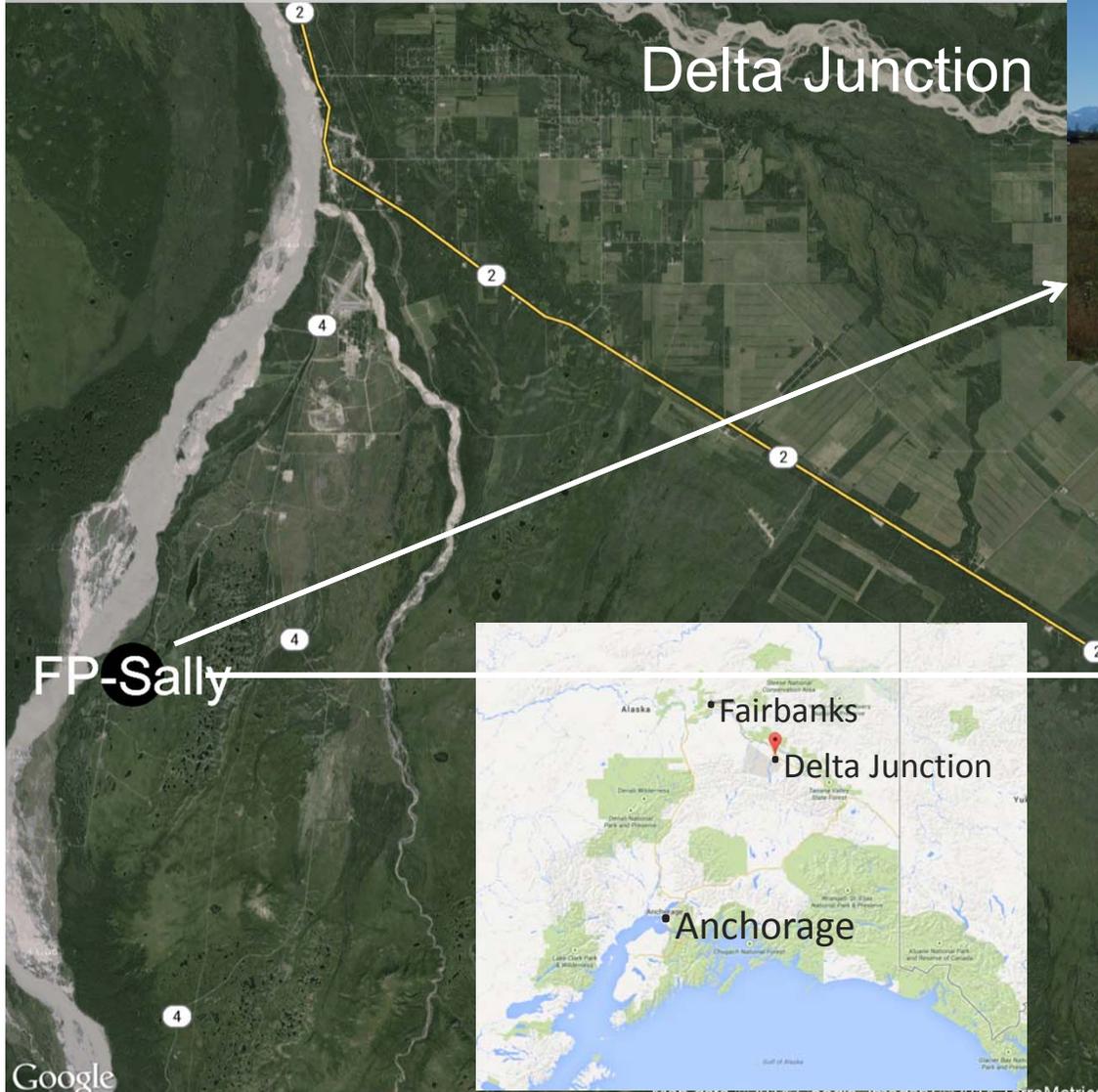
## Donnelly Training Area



Pacific Northwest

ISTORY

Since 1965



Firing Point Sally



Southwest Corner of DU 1

# Contaminants of Concern



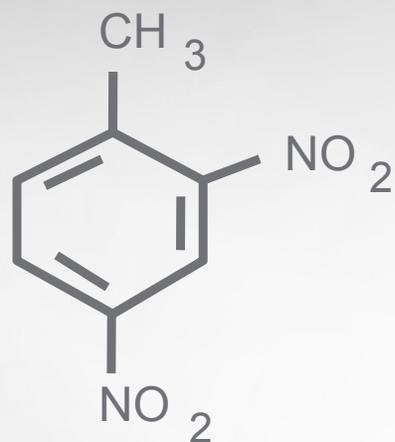
Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965

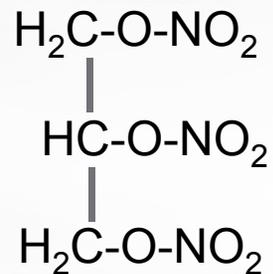
### 105-mm Howitzer



Fibers 0.4 to 7.5 mm long  
containing 2,4-DNT

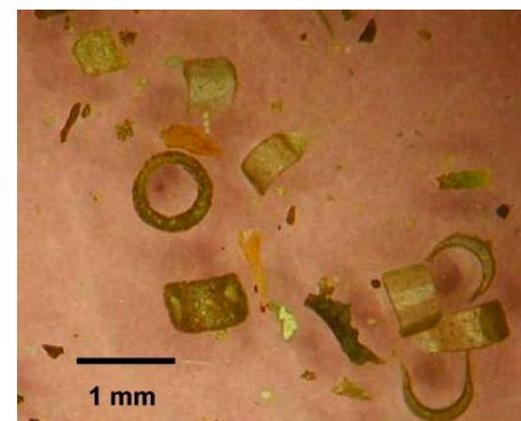


2,4-Dinitrotoluene  
(2,4-DNT)



Nitroglycerin  
(NG)

### 120-mm Mortar



Rings and fragments containing NG

## Solid Residues of Propellants at Firing Points

# Alaskan Training Ranges

## Donnelly Training Area

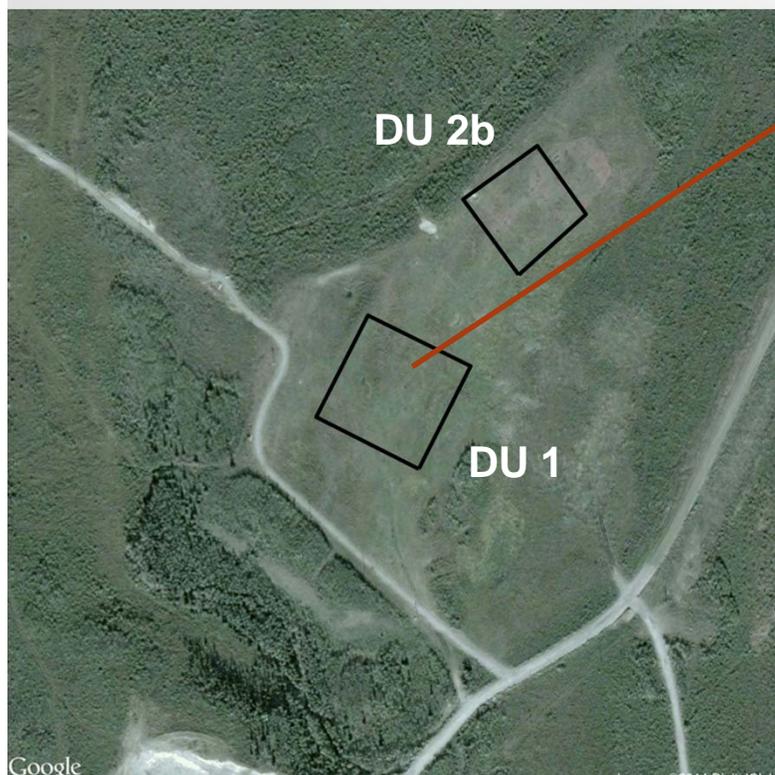


Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965

- DU 1 matches historical footprint
- DU 2b shifted northeast and resized (80 x 80 m.)

- FP Sally chosen due to potential for localized high concentration of 2,4-DNT in DU 1
- Historical ISM samples show some highly variable observations



Month-Year	2,4-DNT	NG
July-2003	1.1, 0.14, 0.45, 3.7, 1.8, 1.6	not detected
June-2005	3.4, 4.7, 3.2	not detected
July-2007	2.1, 1.5, 5.4	0.7, 0.7, 0.8
September-2007	<b>3.3, 2.4, 44</b>	0.9, 0.9, 1.2
May-2008	2.6, 3.9, 3.1	0.8, 3.2, 1.1
September-2008	3.1, 1.3, 2.4	0.8, 0.7, 0.9
July-2009	1.5, 1.2, 1.7	1.0, 0.7, 1.4
September-2012	<b>0.80, 0.75, 8.0</b>	0.7, 0.4, 0.4

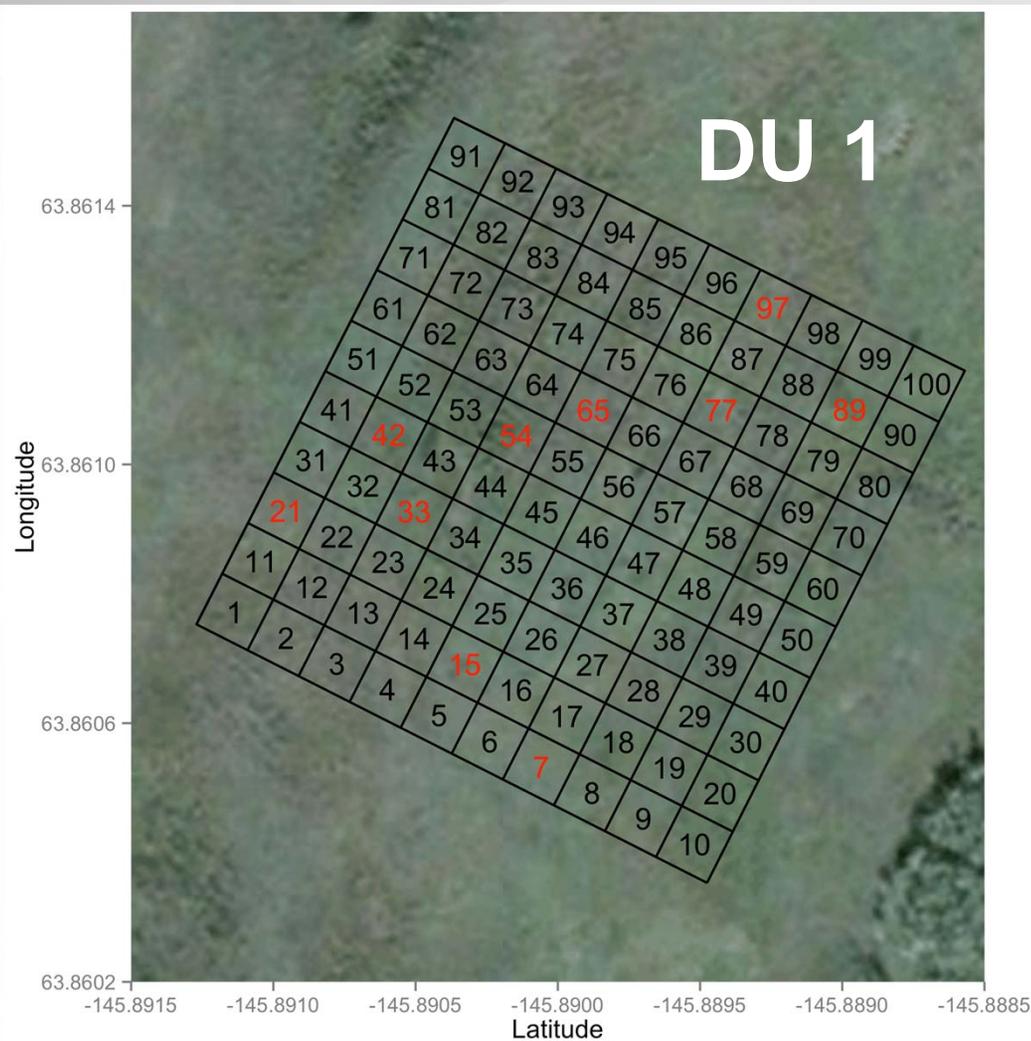
# Donnelly Training Area

## DU-1 FP-Sally



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



- **Super Cells colored red**
- **> 23,000 increments collected**
- **>200 of ISM samples collected**



# Method 8330B Sample Prep



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



# Performance Assessment

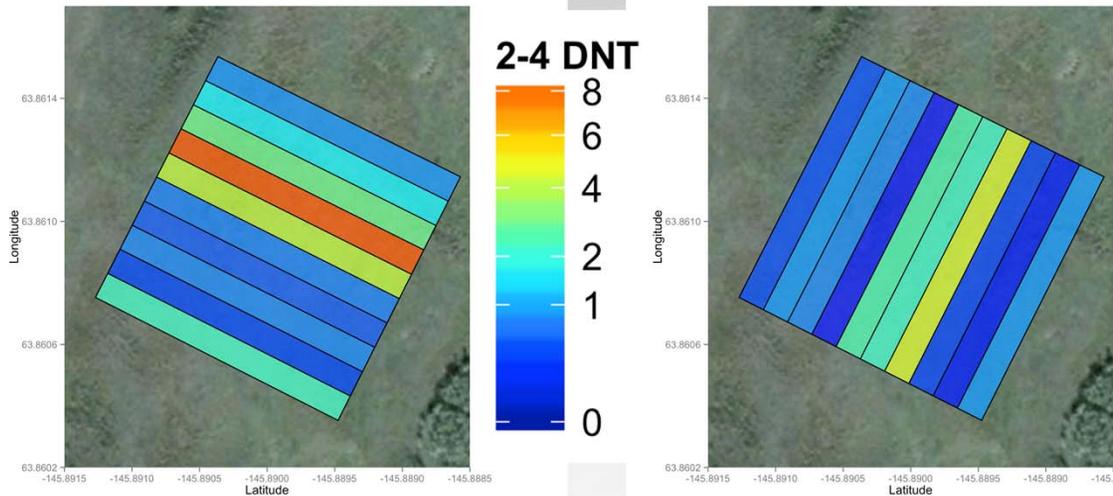


Pacific Northwest  
NATIONAL LABORATORY

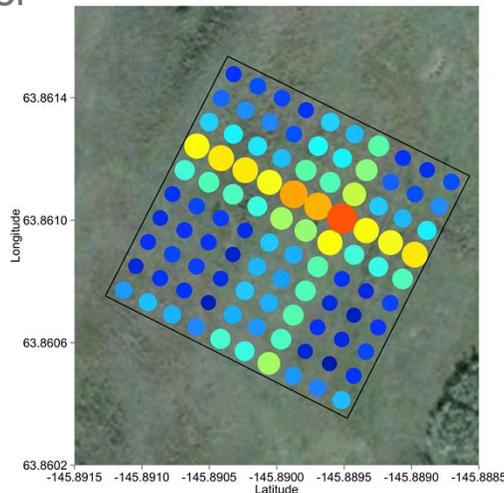
Proudly Operated by **Battelle** Since 1965

$$\frac{(\text{number of columns} \times \text{Row}_i + \text{number of rows} \times \text{Column}_j)}{\text{number of columns} + \text{number of rows}} = \text{Cell}_{ij} \text{ test statistic}$$

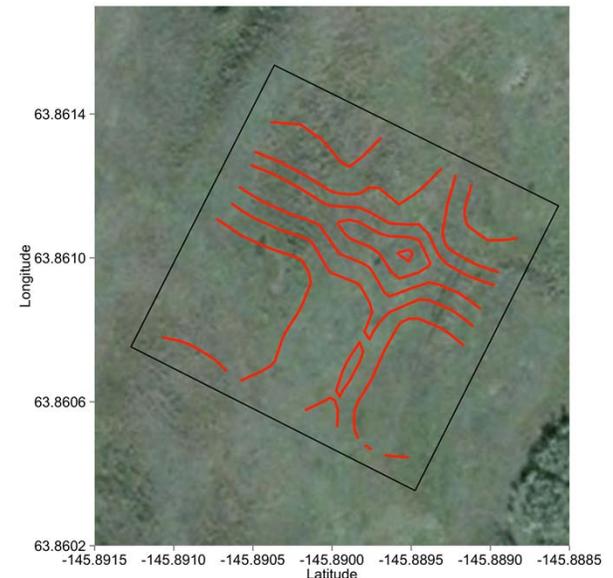
- ▶ Develop elevated cells contour map from row and column ISM



Calculate cell level test statistics



Develop elevated region contour map

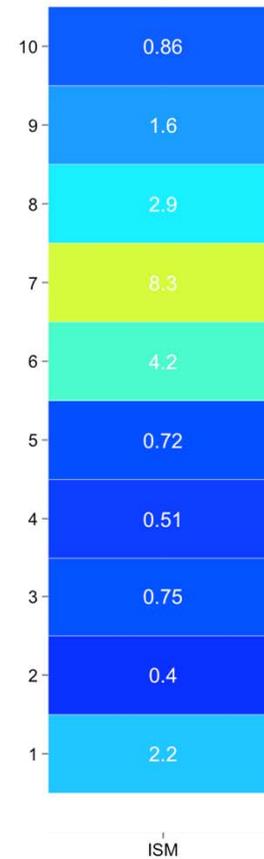
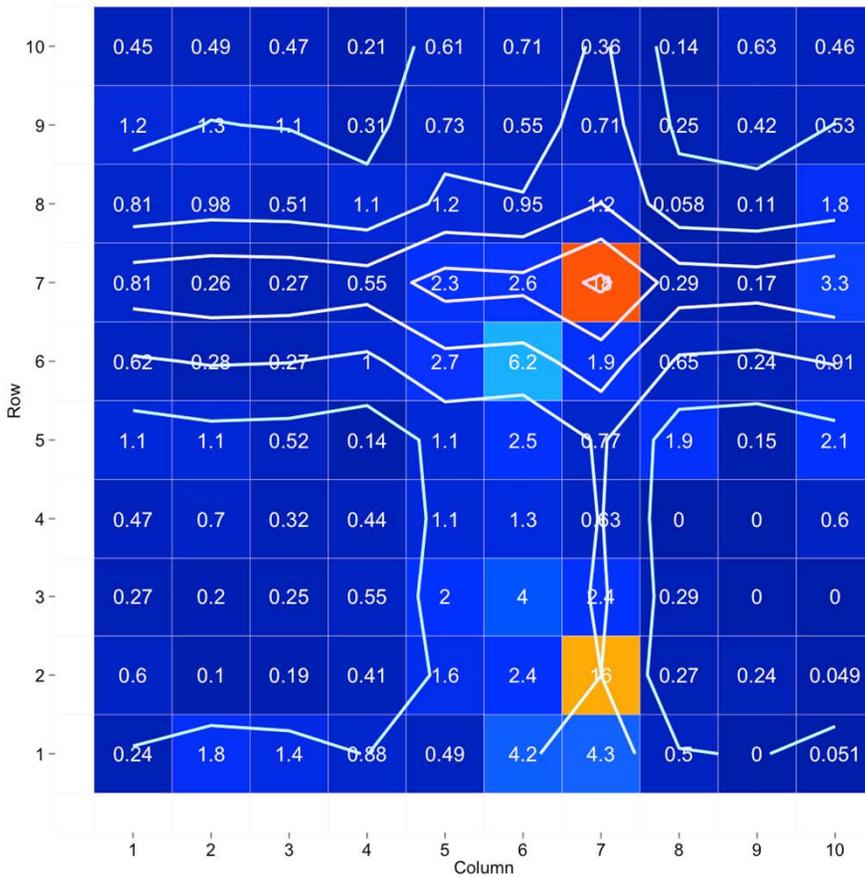


# Performance Assessment



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965

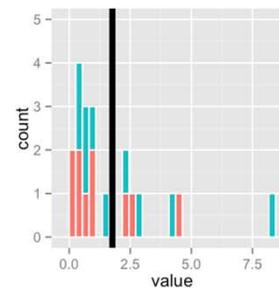
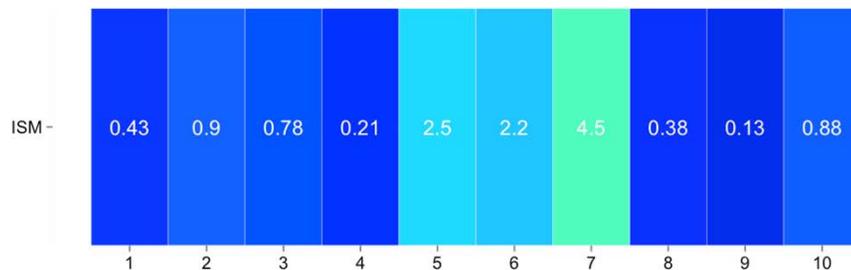


## Row/Column ISM

- Contour estimates
- Mean estimate of 1.8 mg/kg 2-4 DNT

## Cell level ISM

- Colored grid cells
- Mean estimate of 1.3 mg/kg 2-4 DNT



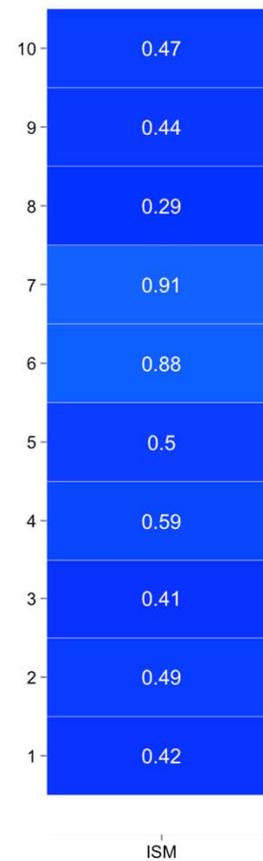
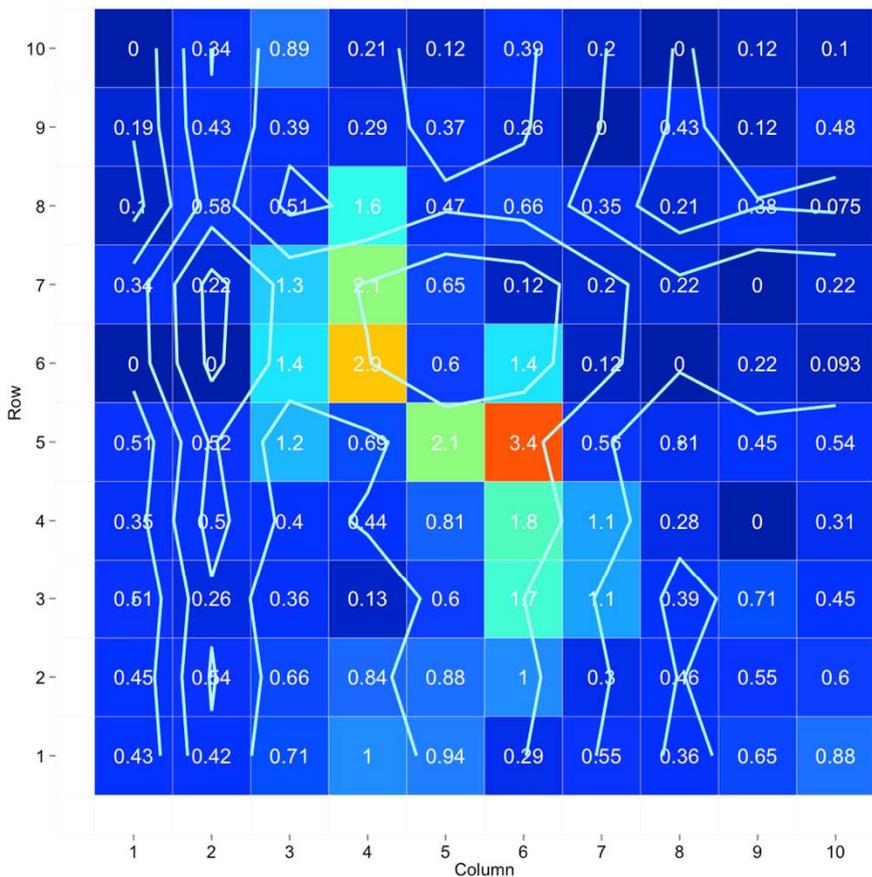
Mean and spatial mapping align

# Performance Assessment

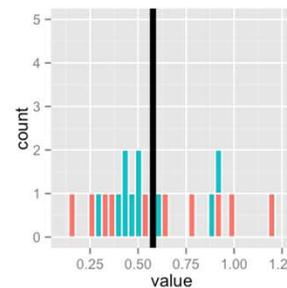
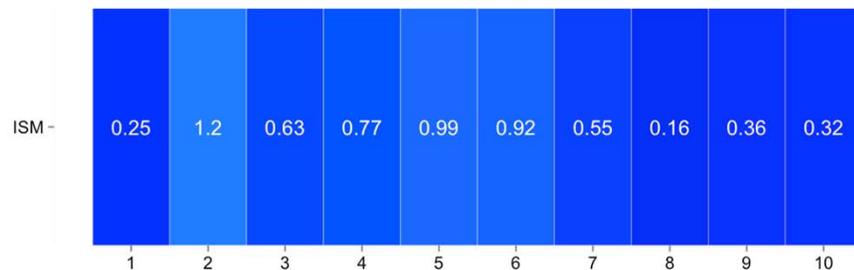


Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



- ▶ **Row/Column ISM**
  - Contour estimates
  - Mean estimate of 0.6 mg/kg NG
- ▶ **Cell level ISM**
  - Colored grid cells
  - Mean estimate of 0.6 mg/kg NG



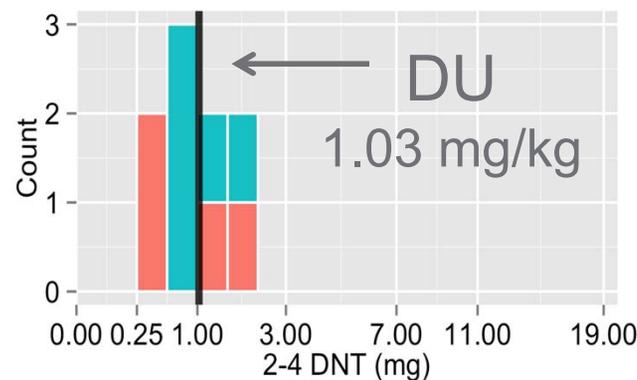
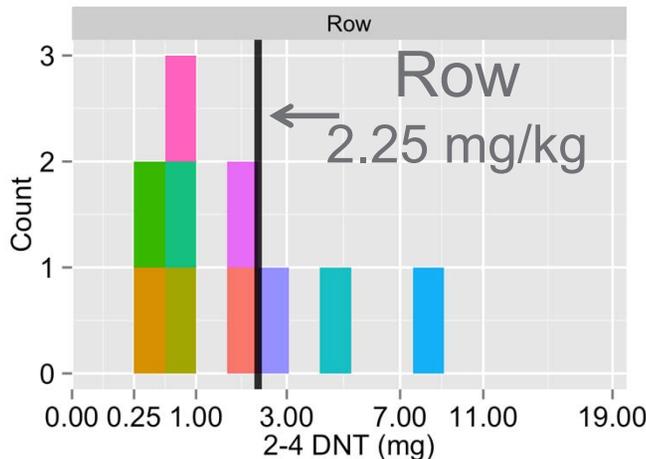
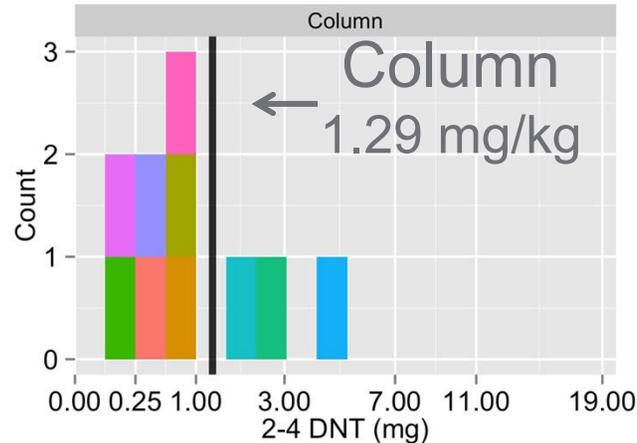
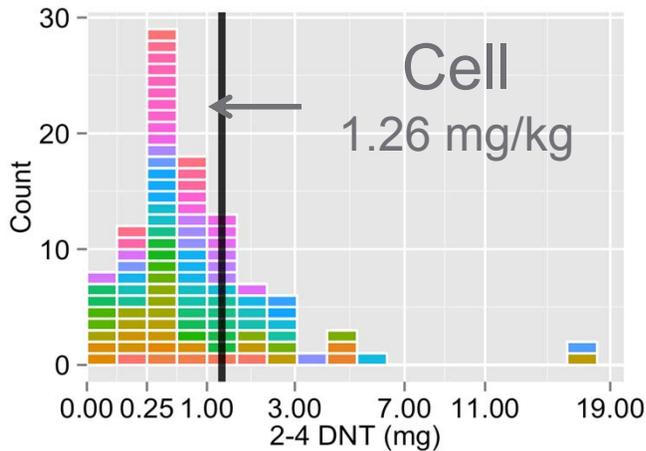
Mean and spatial mapping align

# Performance Assessment



Pacific Northwest  
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



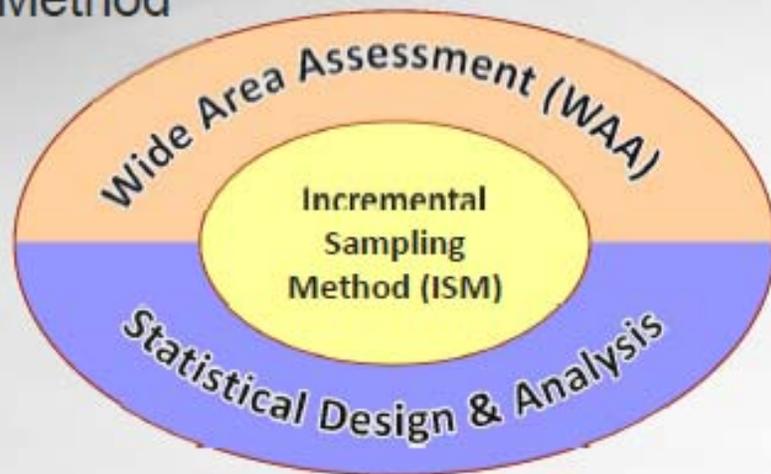
Replicate █ Traditional MIS █ VSP

FP-Sally, DU 1:  
2-4 DNT

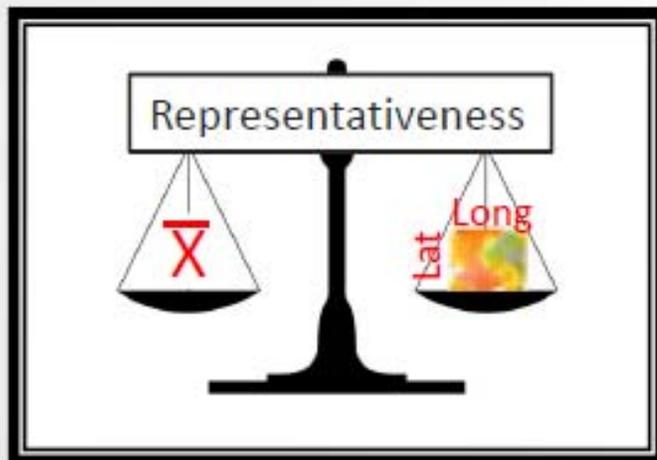
- ▶ **DU mean estimates are consistent across ISM methods**
- ▶ **“True” mean is 1.5 mg/kg for 2-4 DNT**

# Conclusions

- ▶ Statistically based sample designs need to be more tightly incorporated into the Incremental Sampling Method



- ▶ Dialogue between ISM “devotees” and discrete sampling “devotees” needs to continue



- ▶ We need to strike a better balance between spatial extent and mean estimation