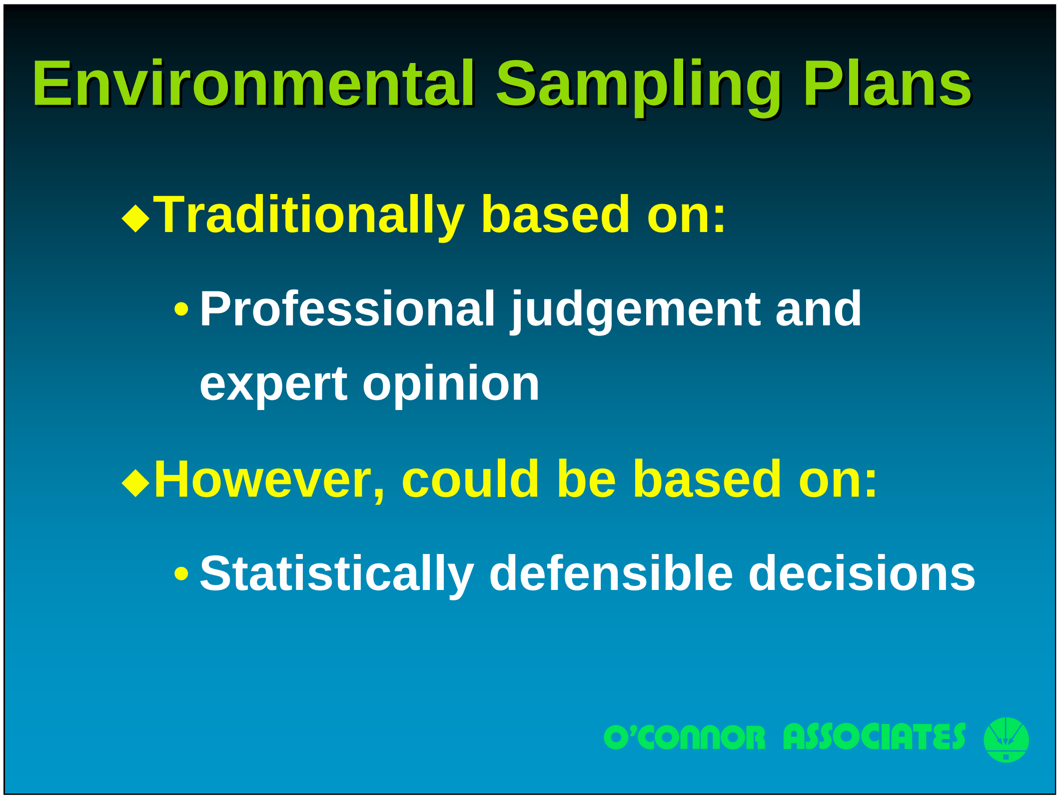
Are You Certain That You've Taken a Sufficient Number of Soil Samples?

An Introduction to Statistically Defensible Sampling Plans

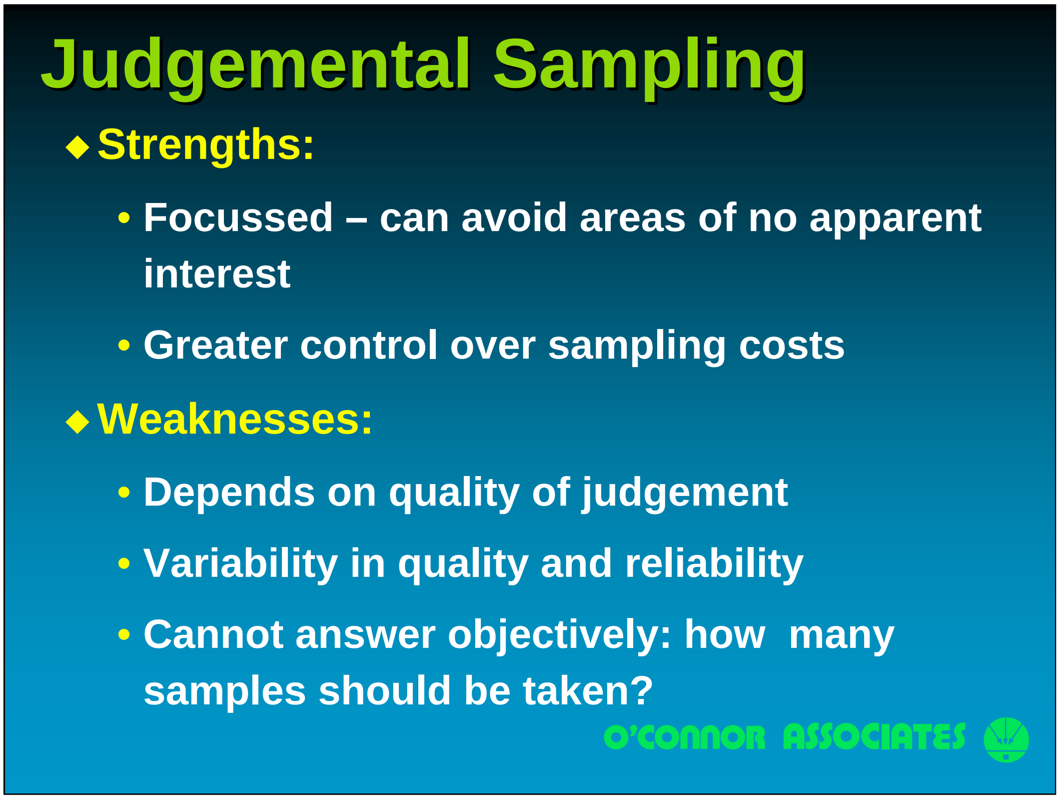
by

Tai Wong, John Agar & James Carss

O'CONNOR ASSOCIATES



Traditionally, environmental sampling plans (soil, groundwater, surface water, etc) have been conducted in an “ad hoc” manner by the environmental assessment team members using expert opinion and professional judgement. Recently, sampling plans based on statistically defensible decisions are gaining popularity especially in the US through the support of the EPA.



Strengths:

-focussed – can avoid areas of no apparent interest and concentrate on areas of greater interest and may be perceived to give better coverage in area of concern or potential concern

-Greater control over sampling costs

Weaknesses:

-subjective – depends on quality of judgement; based on experience, therefore, high variability in quality and reliability; Non-random sampling. Therefore, statistically biased and cannot be defended statistically –
tend to look at "worst case" scenarios

- cannot determine if results met tolerable limits on uncertainty required for decision
- cannot use data to calculate valid unbiased estimate of mean or variance
- cannot make statements about the probability that a hot spot was missed

Statistical Sampling

- Strengths
  - Provides valid inference to population parameters, e.g. mean and variance
  - Can optimize design – balance between uncertainty and cost
- Weaknesses
  - More complicated process
  - More samples => higher costs
  - Cannot handle point source
  - Limited to populations where values are not spatially or temporally correlated

O'CONNOR ASSOCIATES

Statistics based:

-Supports valid inference to the target population- provides sampling basis when judgement cannot

-Supports quantitative estimates of uncertainty and variance

-Provides estimates of limits on uncertainty required for the decision

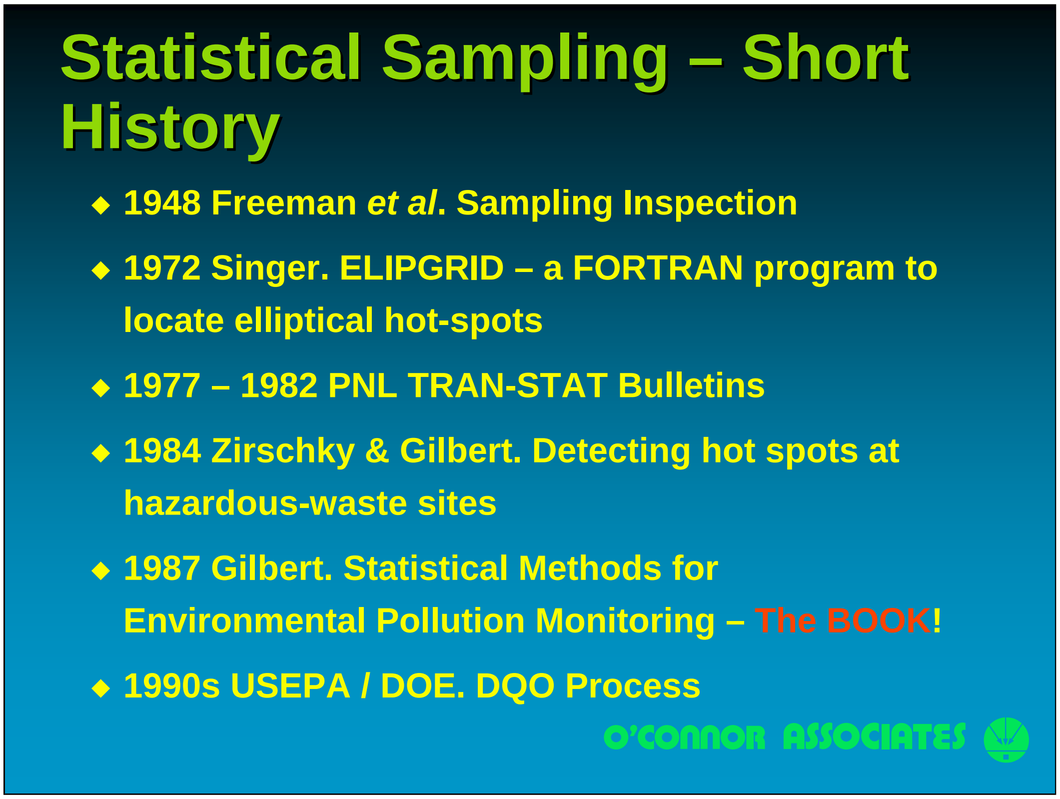
-Randomness provides assurance that anything unexpected in the CSM has not been missed

Weaknesses

-harder to explain to general public – "mathematical" complications

-Chosen sampling locations may appear not to make any sense

-Little perceived control on higher sampling costs



Started off with locating hot-spots and or buried radioactivity from nuclear weapons testing.

Throughout this history, PNNL and Gilbert are very much involved.

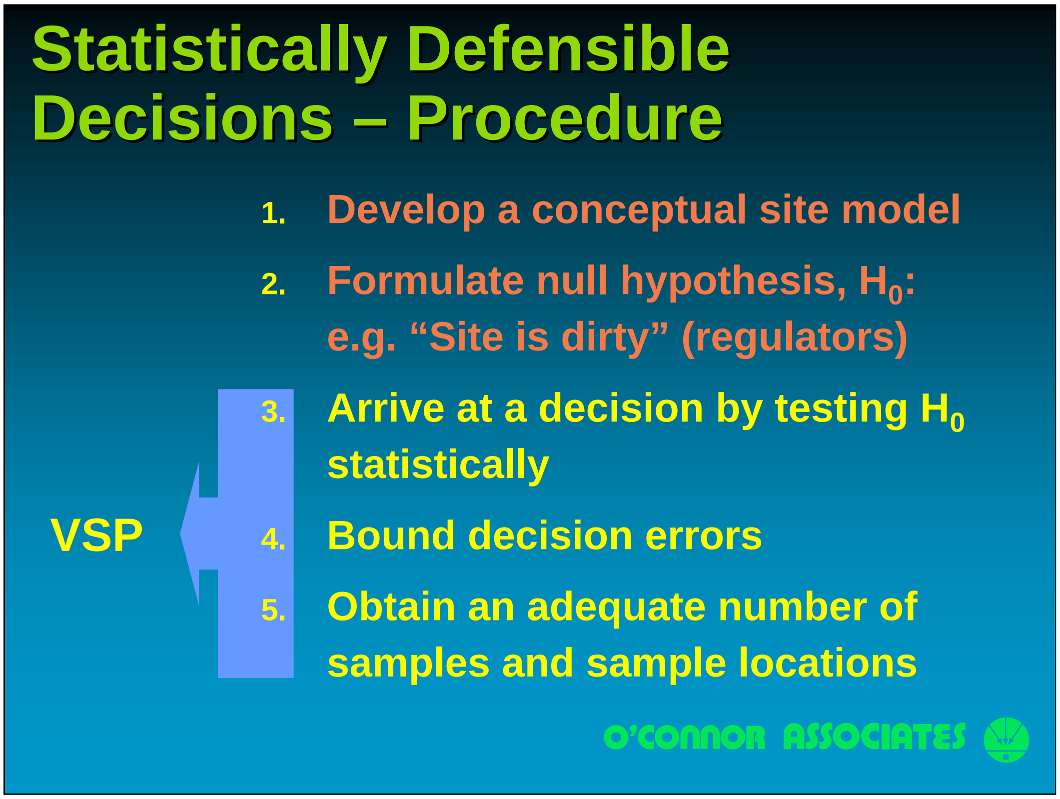
•1948 Freeman *et al*. Sampling Inspection

•1972 Singer. ELIPGRID – a FORTRAN program to locate elliptical hot-spots. Geocom Programs 4:1-16.

•1977 – 1982 Pacific Norwest Laboratory. TRAN-STAT Statistics for Environmental Studies. 19 Bulletins through the years 77 to 82.

•1978 Doctor & Gilbert. Two studies in variability for soil concentrations. Selected Environmental Plutonium Research Reports of the Nevada Applied Ecology Group.

•1984 Zirschky & Gilbert. Detecting hot spots at hazardous-waste sites, Chemical Engineering.



Procedure for arriving at statistically defensible decisions

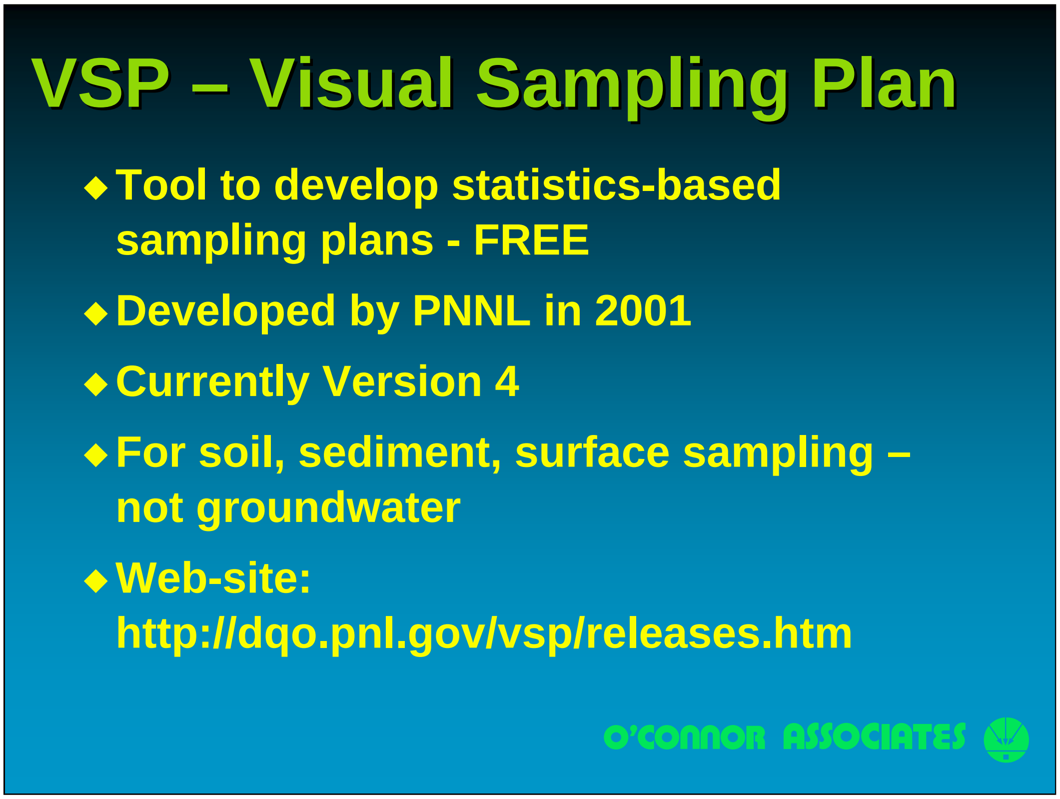
-Based on a site conceptual model - develop a null hypothesis – like a default condition – usually for an ESA - the regulator’s prerogative is “Site is Dirty”

-Arrive at decision by testing Ho statistically – accept or reject

-Use sufficient number of samples

-Establish the limits of decision error

-Items 3 to 5 can be addressed using VSP software



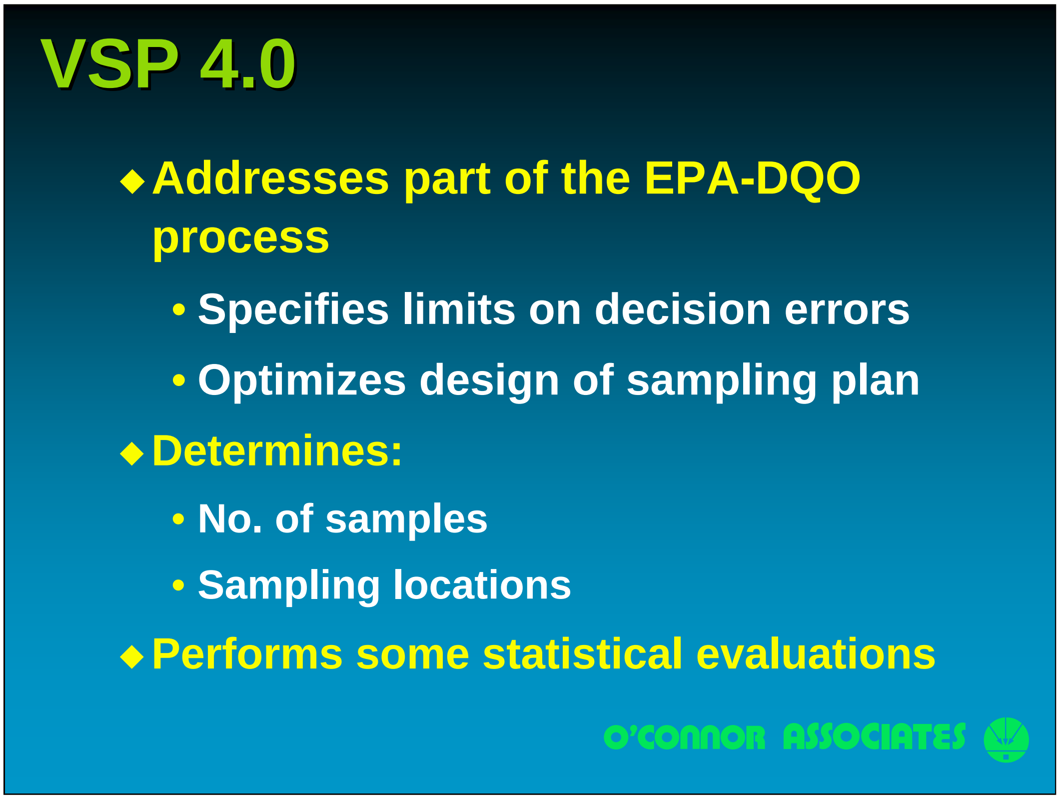
PNNL – Pacific Northwest National Lab

Development has been ongoing since 2001 but the statistical methodologies have been published by Gilbert and others between mid-1970s and the present. “Statistical Methods for Environmental Pollution Monitoring”

Surface sampling – on surface of walls around rooms in a building – e.g. dust

Currently limited to non-point source, e.g. not to be used to chase a g.w. plume

Free – downloadable from the web

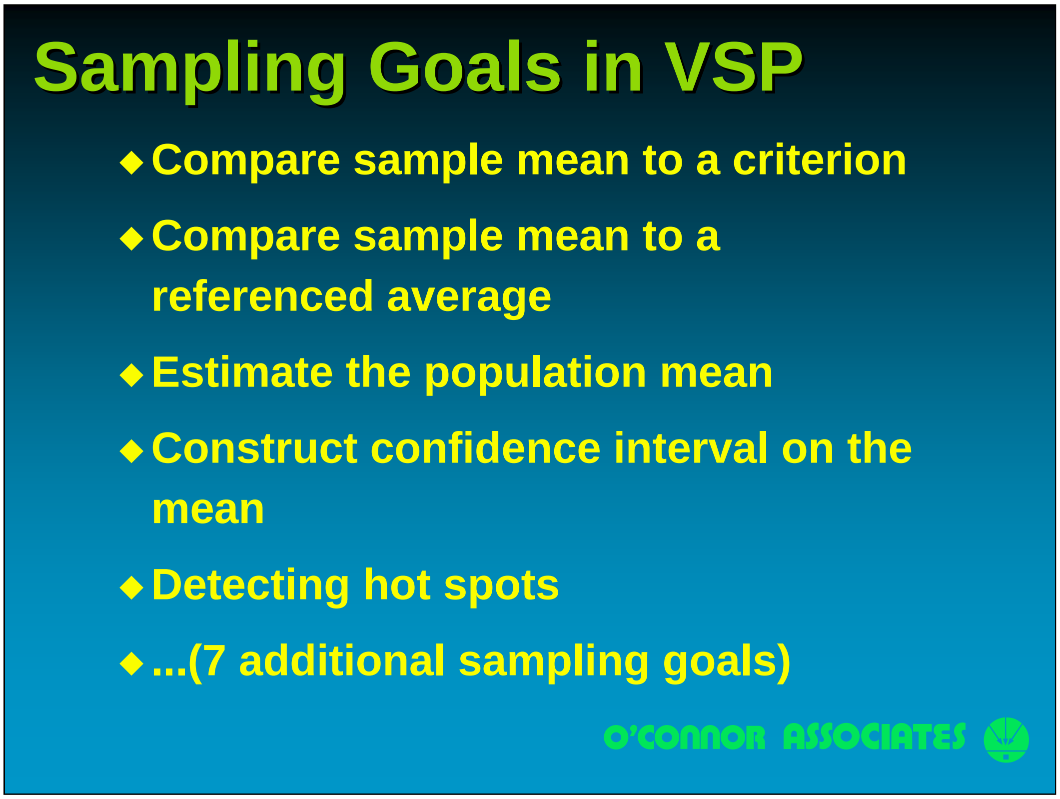


Steps 6 and 7 of the DQO process.

Best illustrated using some examples.

DQO Process:

1. State the problem
2. Identify the decision
3. Identify input to the decision rule
4. Define the boundary of the problem
5. Develop a decision rule
6. Specify the tolerable limits on decision errors
7. Optimize the design for obtaining data



Common Sampling Goals in VSP – 5 listed – altogether 12

1. Compare Average to Fixed Threshold: calculates number of samples needed to compare a sample mean or median against a predetermined criterion and places them on the map.
2. Compare Average to Reference Average: calculates number of samples needed to compare a sample mean or median against a reference mean or median and places them on the map.
3. Estimate the Mean: calculates number of samples needed to estimate the population mean and places them on the map.
4. Construct Confidence Interval on Mean: calculates number of samples needed to find a confidence interval on a mean and places them on the map.
5. Locating a Hot Spot: use systematic grid sampling to locate a Hot Spot.

The other 7 are:

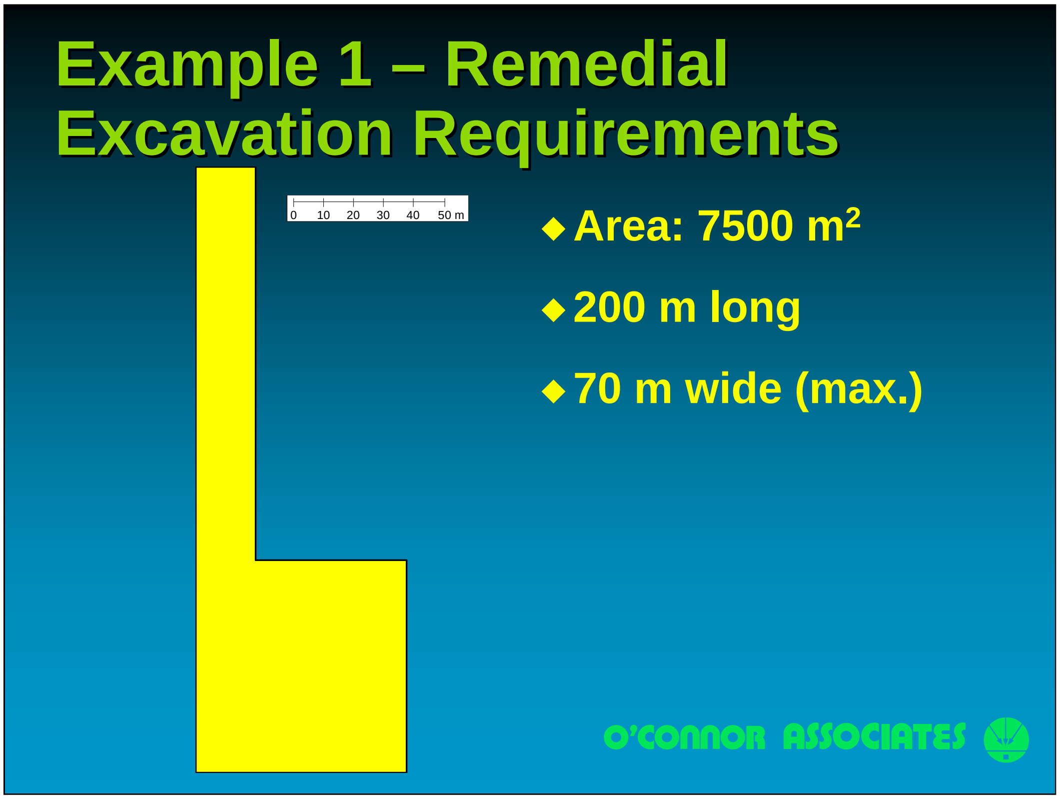
1. Compare Proportion to Fixed Threshold: calculates number of samples needed to compare a proportion to a given proportion and places them on the map.
2. Compare Proportion to Reference Proportion: calculates number of samples needed to compare two proportions and places them on the map.
3. Estimate the Proportion: calculates number of samples needed to estimate the population proportion and places them on the map.
4. Find UXO Target Areas: traverse and detect an elliptical target zone using swath sampling.
5. Assess Degree of Confidence in UXO Presence.

How to Use VSP?

- **Two examples to illustrate developing a sampling plan:**

1. **Sampling a study area for remedial excavation requirements**
2. **Sampling remedial excavation walls to determine probable "hot spot" size**

O'CONNOR ASSOCIATES

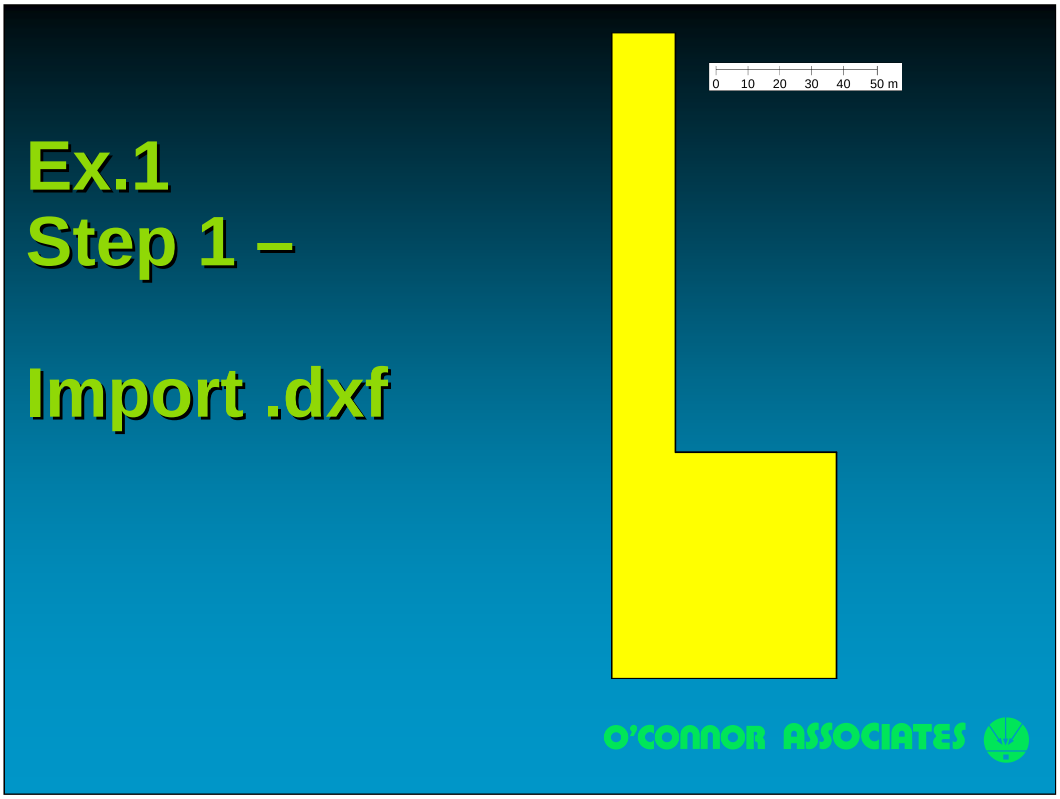


Want to sample the area "study area" for metal impacts to determine remedial excavation requirements with a high level of confidence, e.g. >90%.

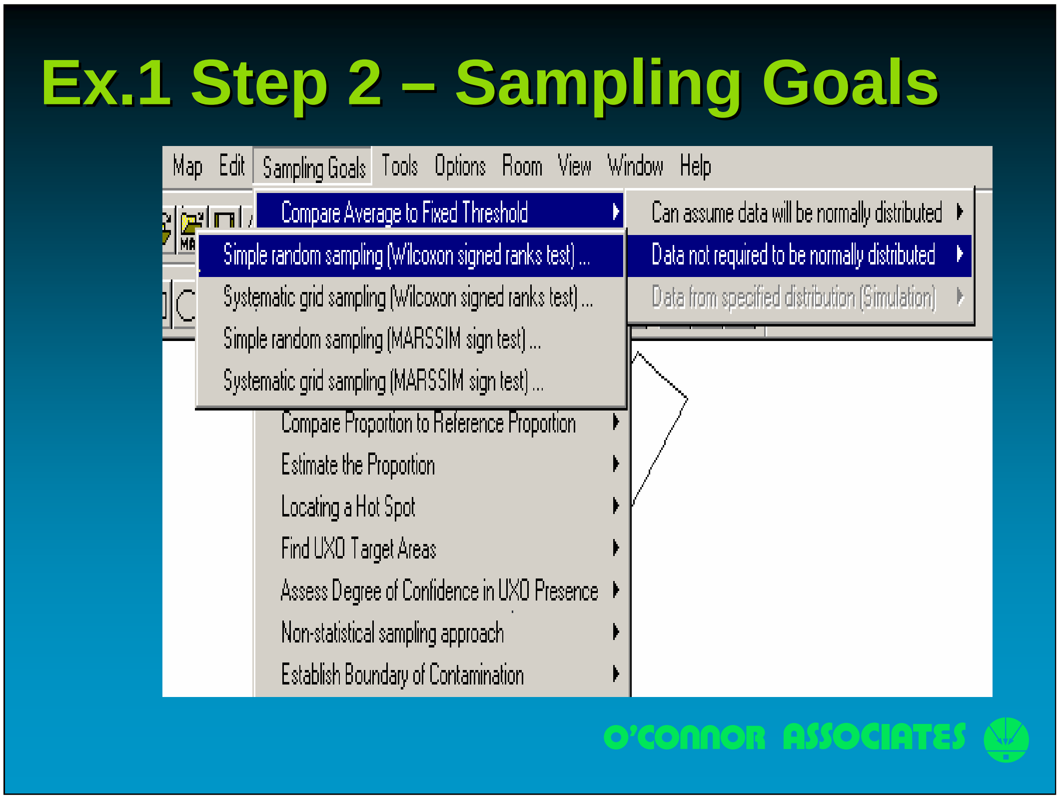
Example 1 - Procedure

1. Import/Draw map
2. Determine sampling goals
3. Establish DPGD (decision performance goal diagram)
4. Obtain sampling grid

O'CONNOR ASSOCIATES



.dxf seems to be the simplest, although VSP does allow .shp files.

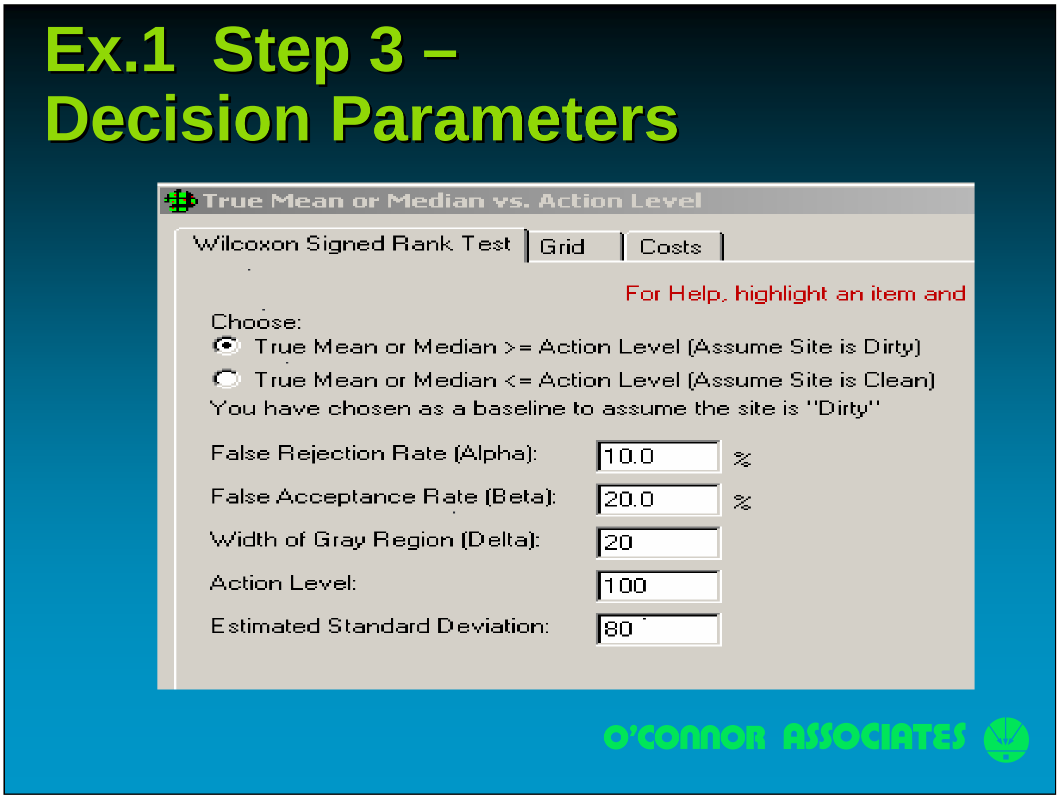


Sampling Goal:

-Compare average to a threshold (site criterion)

-Assume data not normally distributed

-Systematic Grid Sampling – need to use "random" start to be "unbiased"



False Rejection Rate – Type I Error – alpha

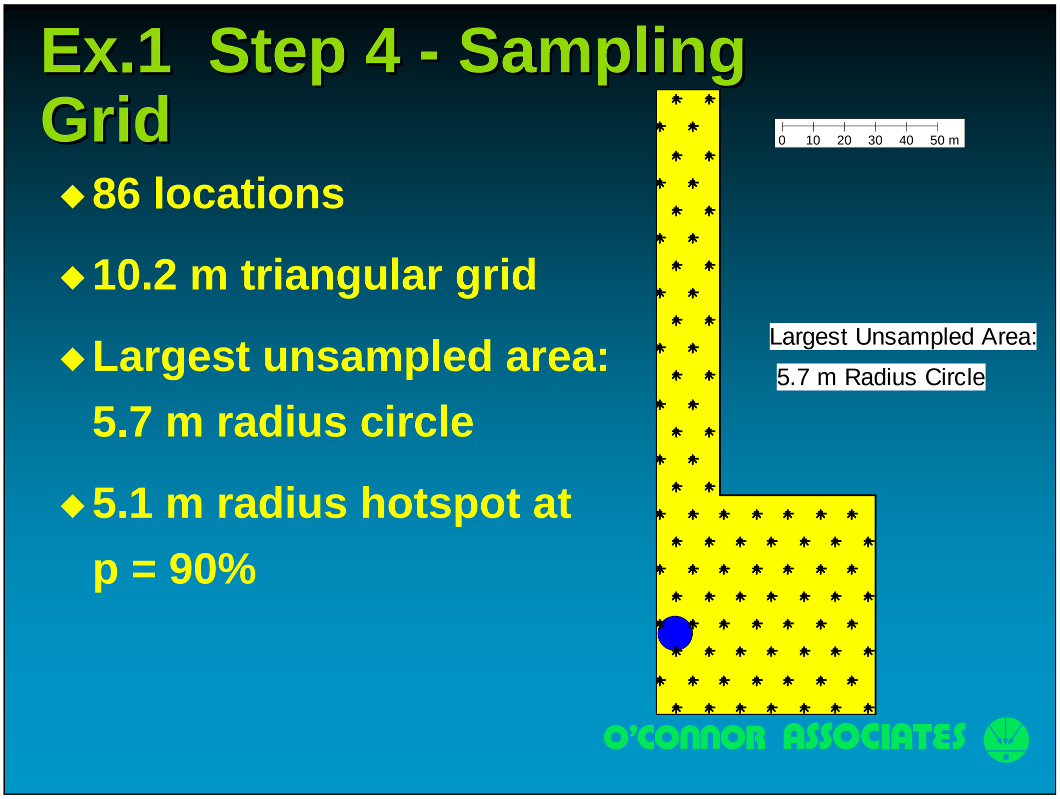
- rejecting H0 when it is true – declaring site is clean when it is actually dirty

False Acceptance Rate – Type II Error – beta

- accepting H0 when it is false – declaring site is dirty when the site is actually clean

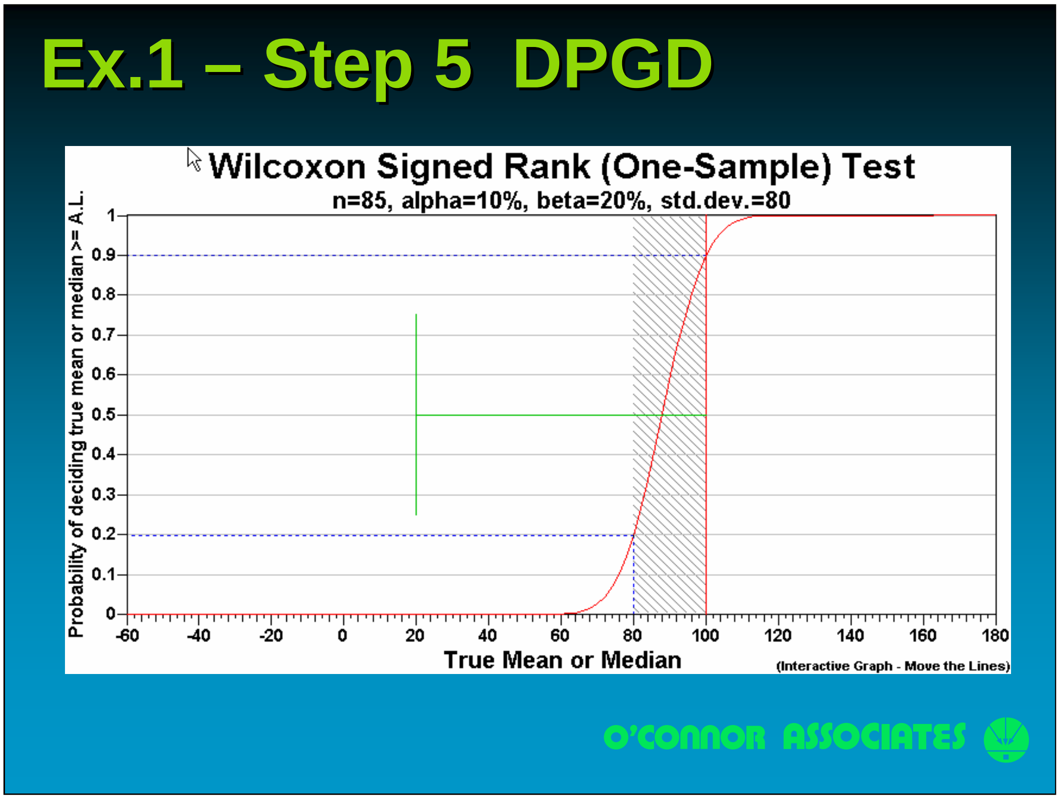
Action Level – equivalent to criteria

Need to estimate standard deviation – by preliminary testing – that means we need to go back and check the grid design after we got the analysis results.



The unsampled area is analogous to the max. size of a "hot spot" that may be missed by this sampling grid;

Approx. 130 m2.



Decision Performance Goal Diagram

α, chance of not detecting locations of contamination and falsely declaring the site is clean when it is actually dirty; we used 10% in this example.

β, chance of sampling a disproportionate number of contaminated locations and declaring the site is dirty when it is actually clean; we used 20% in this example.

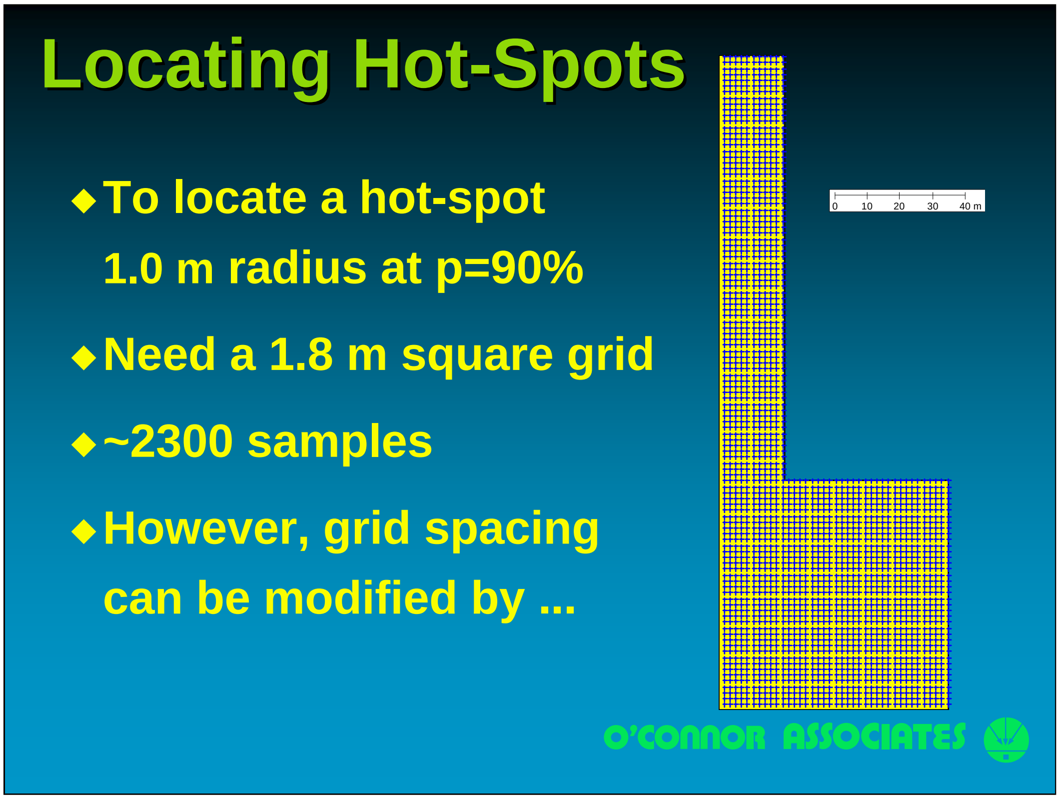
Gray Region – area of uncertainty – between α and β; want to minimize

Number of Samples

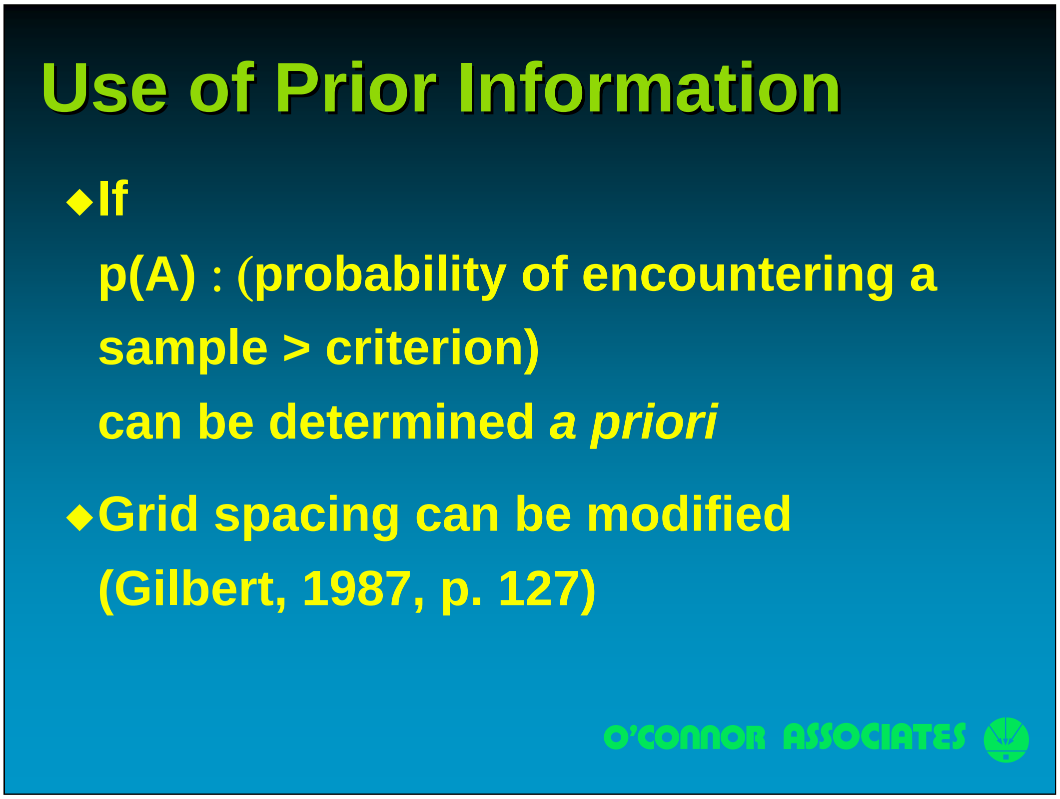
- Based on a symmetric (not necessarily normal) population
- Changes with α, β and s.d.
  - Proportional to $(s.d.)^2$
- Inversely proportional to width of gray region (δ) – uncertainty of the threshold value

O'CONNOR ASSOCIATES

If we keep a=10%, b=20%, and reduce the width of gray region from 20 to 10, N increases from 85 to 336.



Why square grid? – easier to refine by half.



- If probability of encountering a sample > criterion can be determined *a priori.* for example, by pilot testing using a random grid

- Based on theorem of conditional probability

Use of Prior Information (cont'd.)

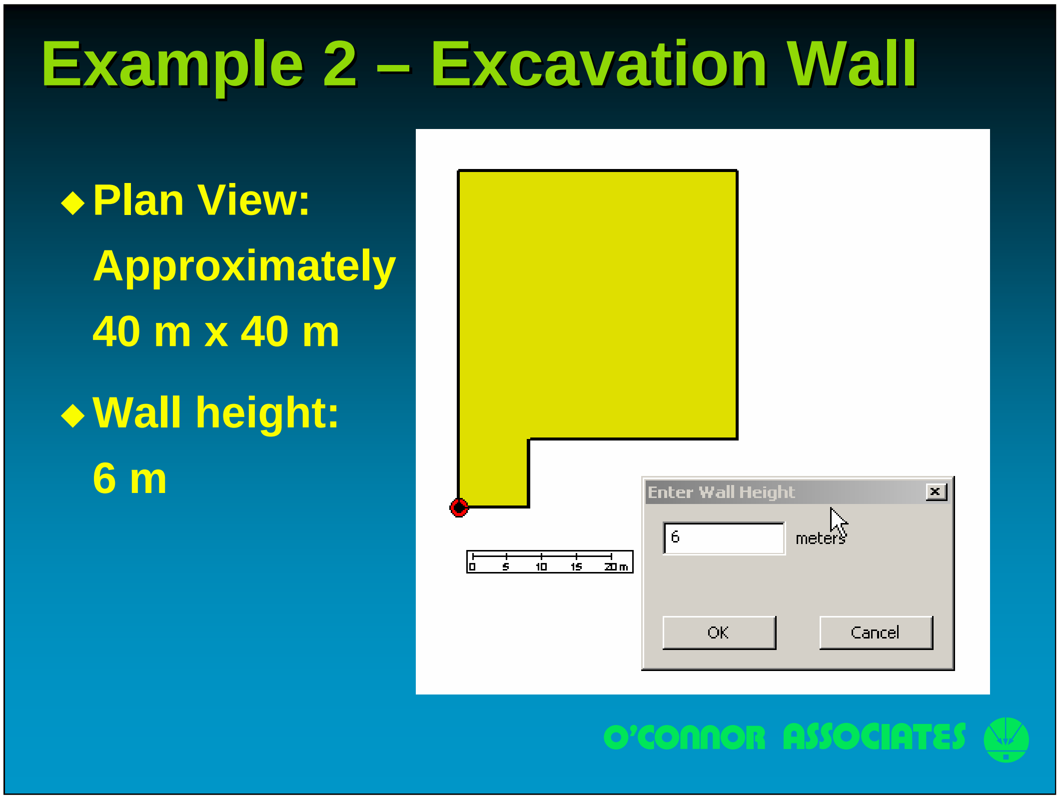
- To locate 1.0 m radius hot spots with 90% probability:

P(A) =	10%	15%	20%	50%
Grid Spacing (m)	10.5	3.0	2.4	1.8
No. Samples	73	835	1302	2300

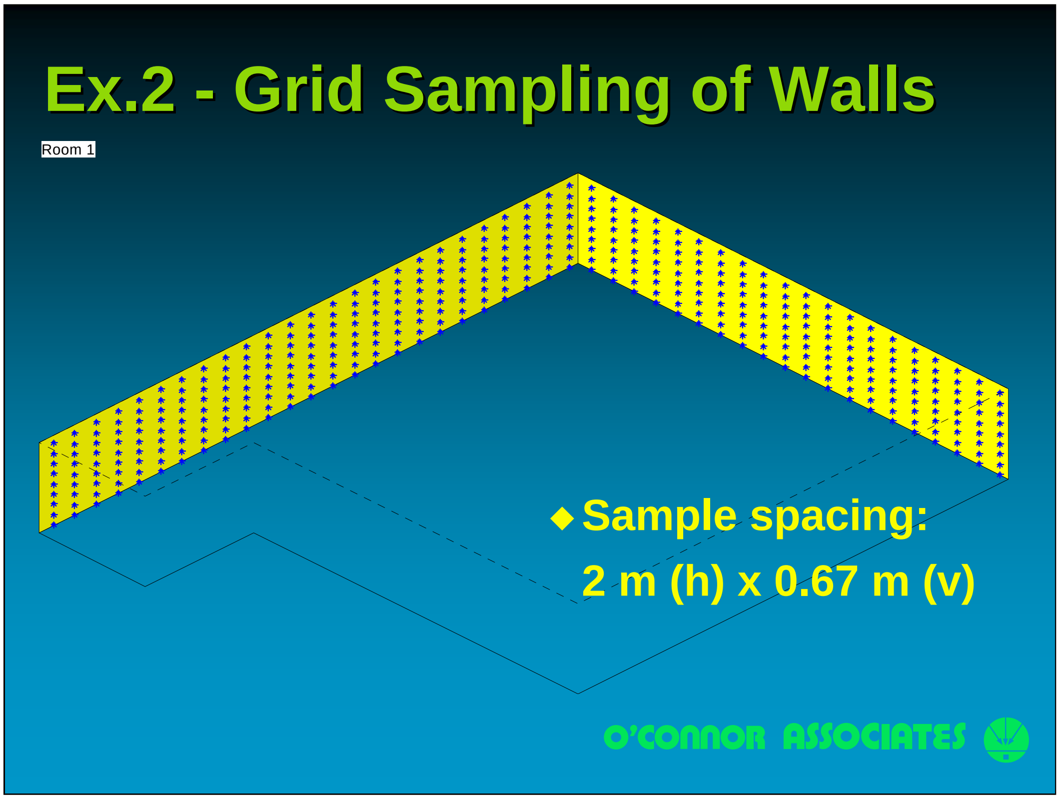
O'CONNOR ASSOCIATES

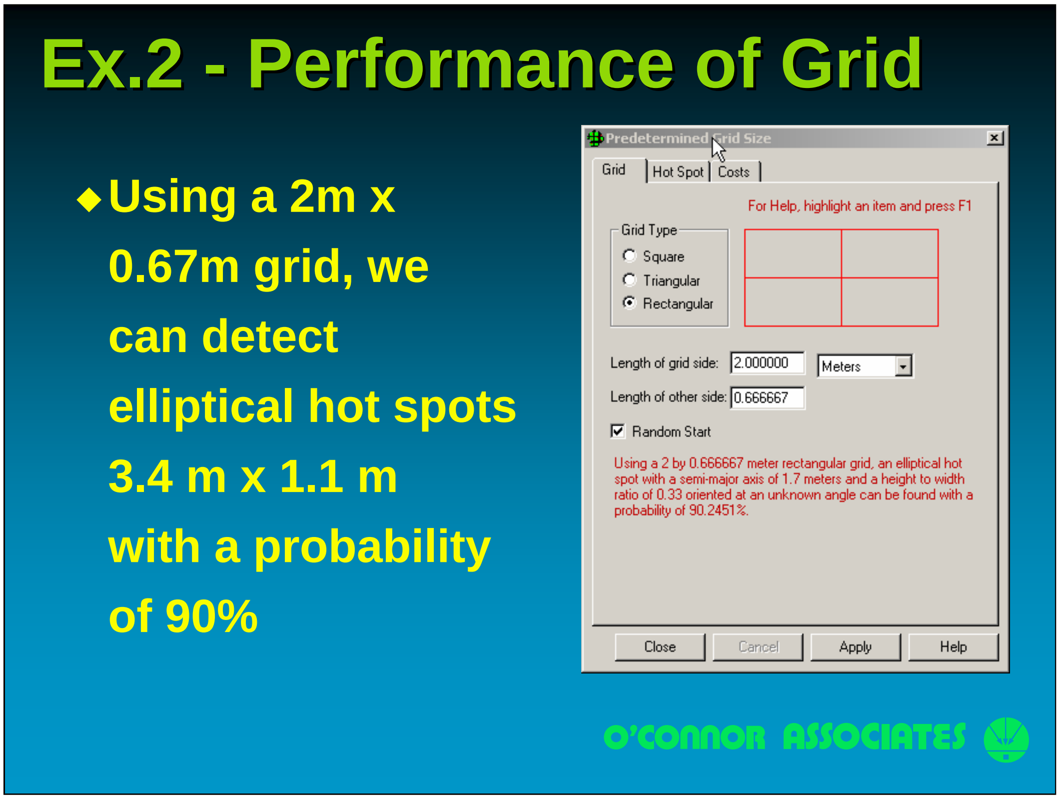
- If p(A) = probability of encountering a sample > criterion can be determined *a priori.* for example, by pilot testing using a random grid

When p(A) is >50%, the original grid size (1.8m) is required.



"Common" problem – sampling on remedial excavation walls to confirm effective removal of the contaminated soils.


Ex.2 - Grid Sampling of Walls
Room 1
Sample spacing:
2 m (h) x 0.67 m (v)
O'CONNOR ASSOCIATES

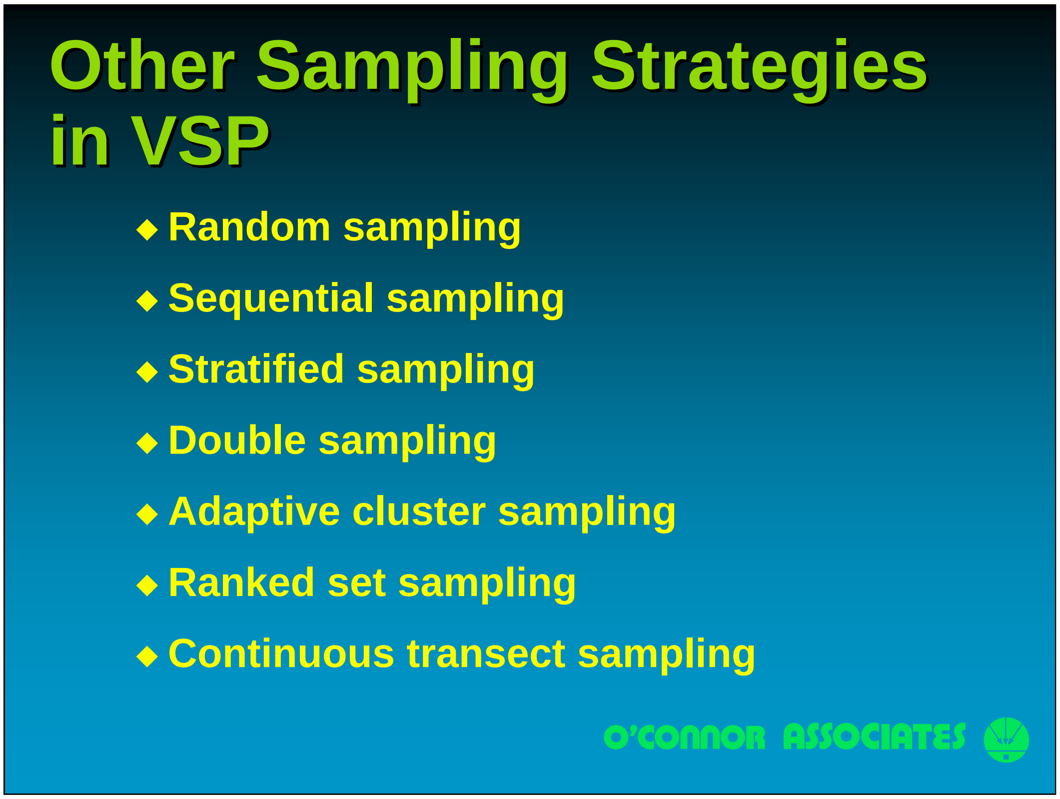


2m x 0.67m - standard OAEI sampling practice.

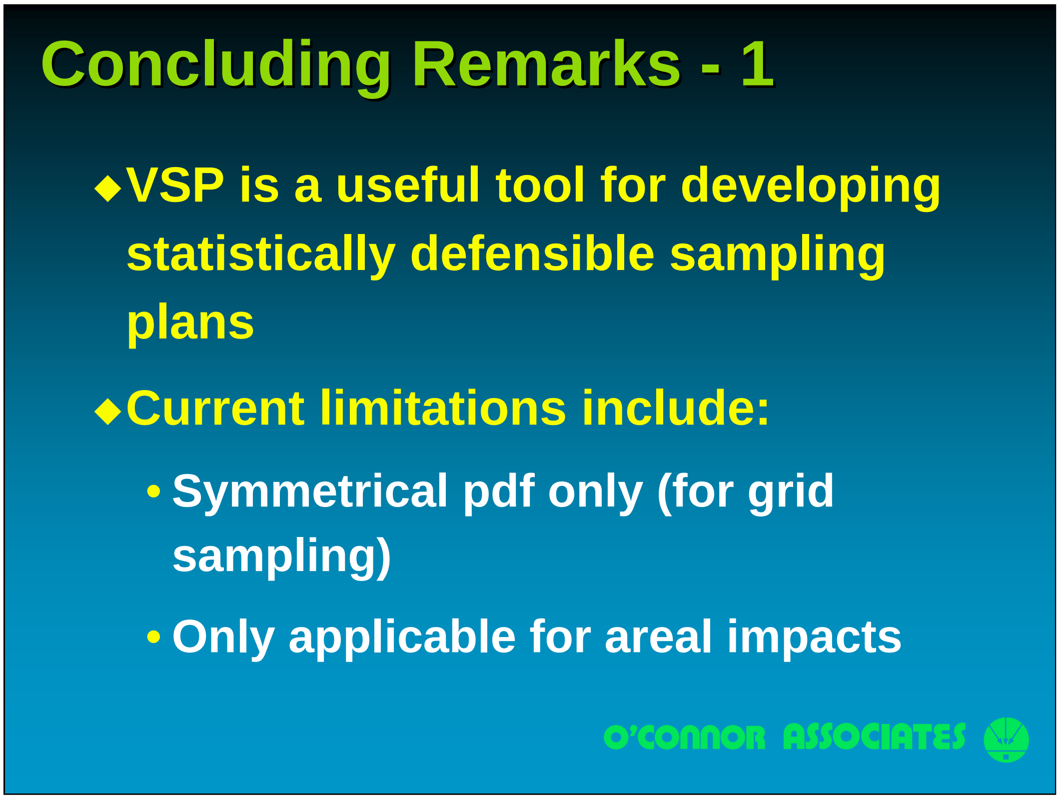
Hotspot size is the maximum size that can be detected.

Actual numbers: 3.4 m x 1.12 m at 90.25%.

Regulators have not specified any sampling requirements.



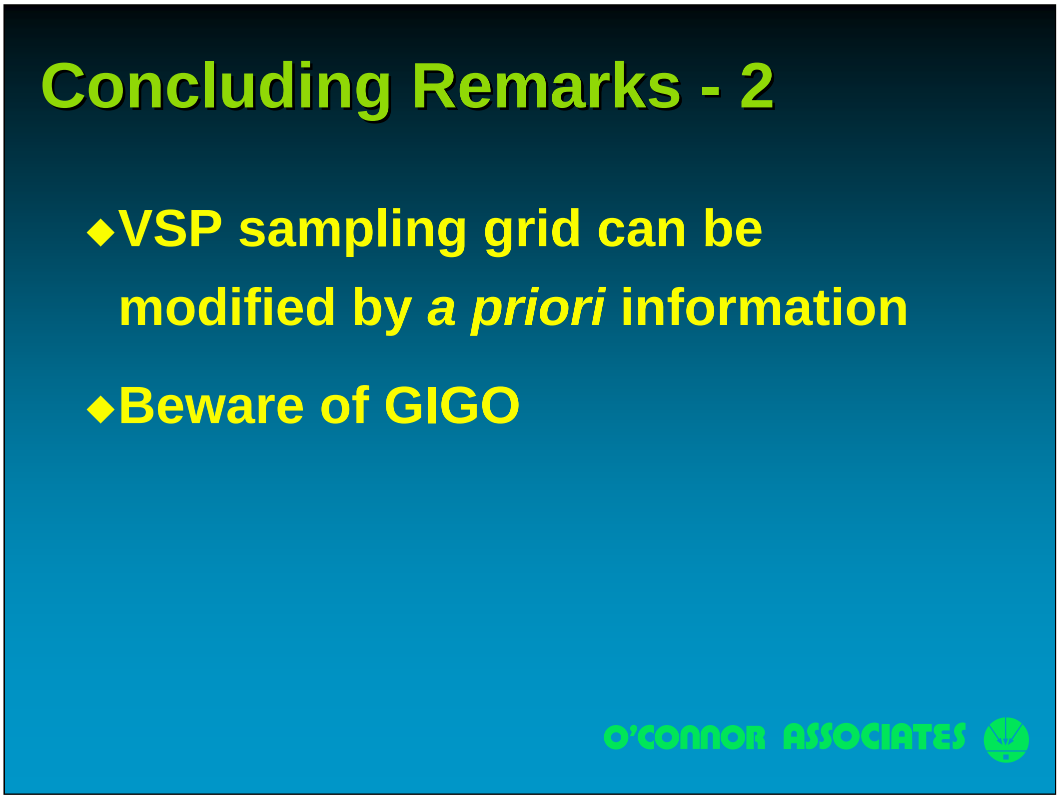
Both examples have used grid sampling, additional sampling strategies available include -



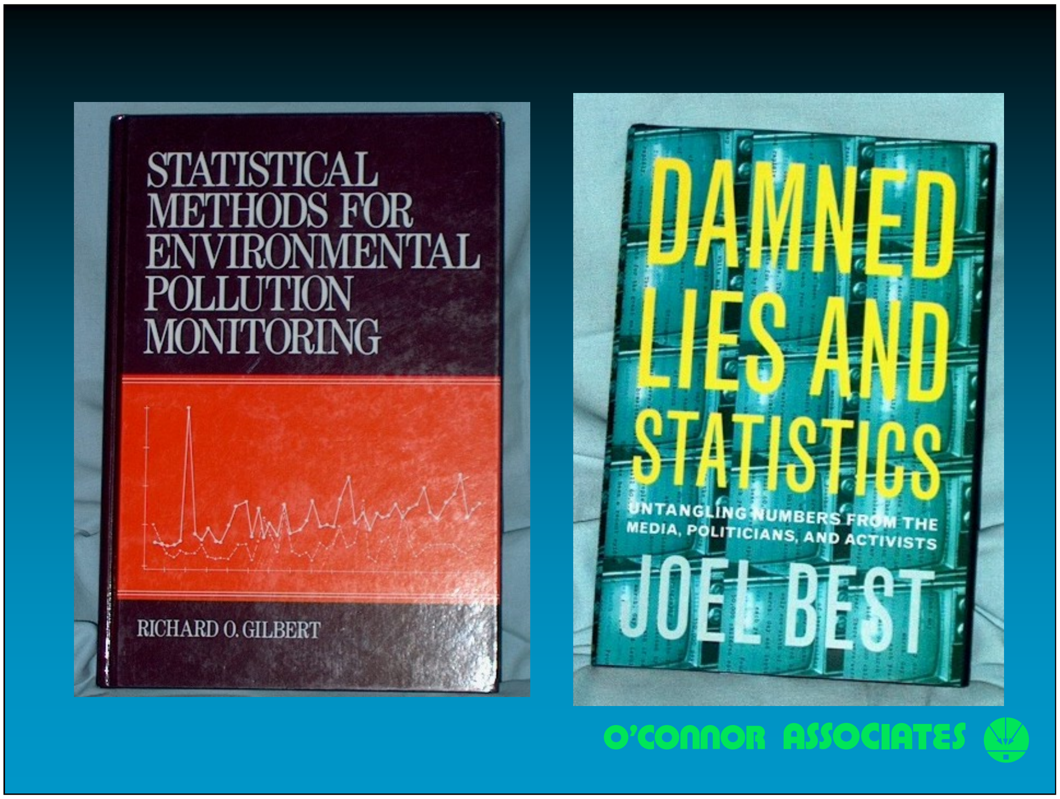
VSP – relatively easy to use and can provide a statistically defensible sampling plan

Symmetrical distribution – can be “circumvented” – CLT, data transform (log data)

c



- Use of a prior information, not built into VSP


STATISTICAL METHODS FOR ENVIRONMENTAL POLLUTION MONITORING
RICHARD O. GILBERT
DAMNED LIES AND STATISTICS
UNTANGLING NUMBERS FROM THE MEDIA, POLITICIANS, AND ACTIVISTS
JOEL BEST
O'CONNOR ASSOCIATES


Thank you
for your attention!
O'CONNOR ASSOCIATES