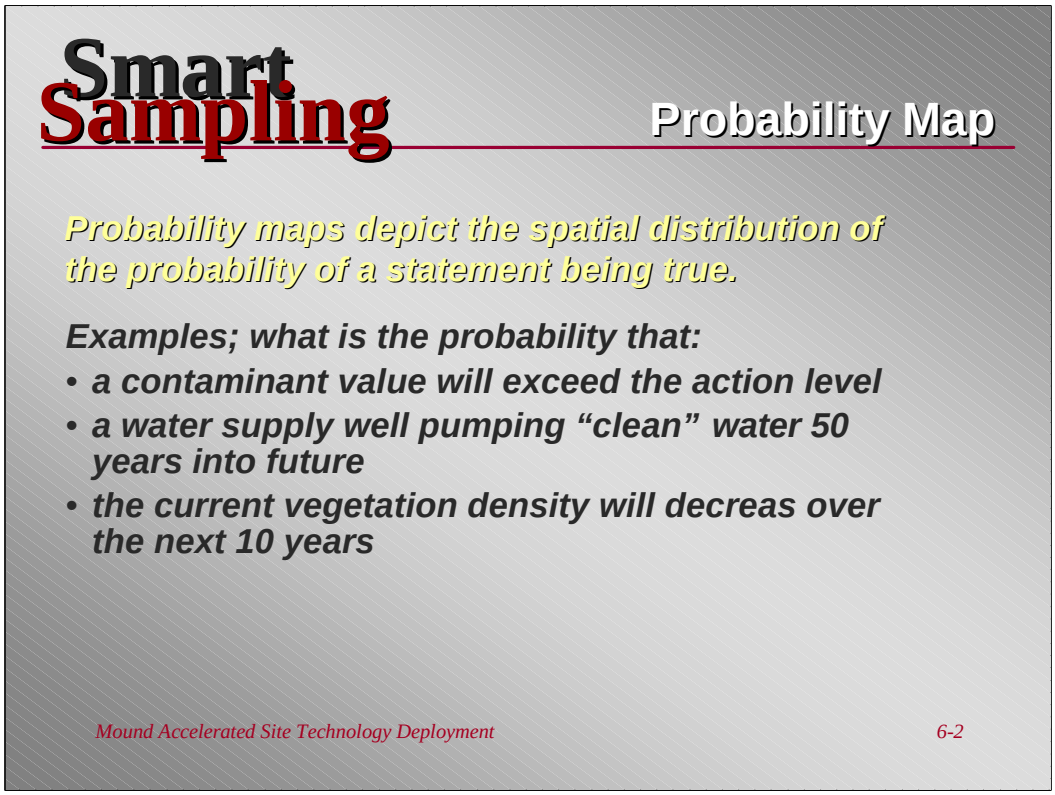
Smart Sampling

Section 6

Probability Mapping



Simulation is a way to generate multiple, equally probable images of what the concentration may look like across a site. We don't want to deal with 100 or 1000 different maps when it comes to cleaning up a site. We want a single remediation map. Building a probability map is the next step.

Smart Sampling

Creating Probability Maps

- ***Basic probability of exceedence maps can be created in two ways:***
  1) ***Transform data (set to 1's and 0's), and then use "kriging" to produce map.***
  2) ***Create multiple realizations of property through "simulation", then determine the empirical probability from distribution at each location.***
- ***Theoretically, the two techniques should give the same map if enough realizations are created.***

  (Practically, implementation of the processes presents so many options, the results usually aren't quite the same)

Indicator Kriging

If the concentration exceeds the threshold at a location, that location becomes a 1; if it does not exceed the threshold at the location, it becomes a 0.

$$z(i) = \begin{cases} 0 < z_k \\ 1 \geq z_k \end{cases} \quad \text{where } z_k = \textit{threshold}$$

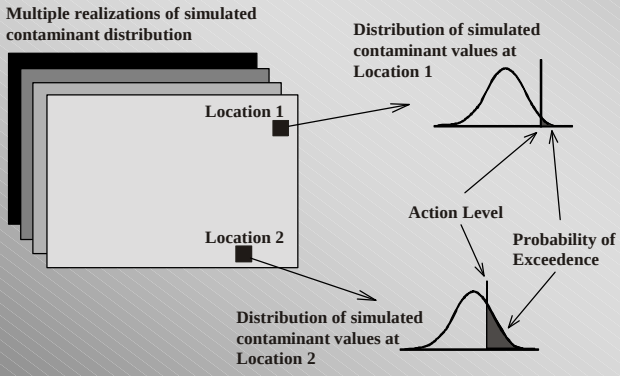
Consider the 0's and 1's to be actual data points and do regular kriging on them to get values between 0 and 1, the probabilities of the statement being true.

Indicator kriging concerns only whether the values are 0 or 1 with relation to the thresholds. Probability indicator kriging is not interested in whether it is 0 or 1 with respect to that threshold, but how close it is.

Smart Sampling

Probability Map

Simulation Post-Processing



The "probability of exceedence" values generated by simulation post-processing are used to create a probability map at a selected threshold.

Mound Accelerated Site Technology Deployment *6-4*

In this approach, you use simulation to create multiple realizations of the property and then create determine empirical probability from distribution at each location.

Smart Sampling

Why Simulation?

SmartSampling has elected to use simulation, even though it is more computationally intensive, because:

- ***Modeling of one variogram is sufficient to solve for multiple thresholds/action levels***

  (If a single threshold has been established for the problem, indicator kriging would probably be the way to go)

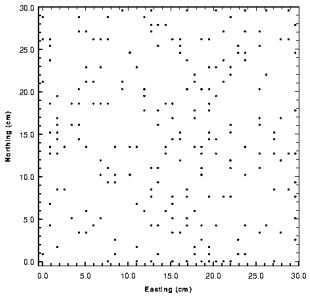
- ***Simulation can be used to address groundwater problems***

  (kriging does not work for groundwater)

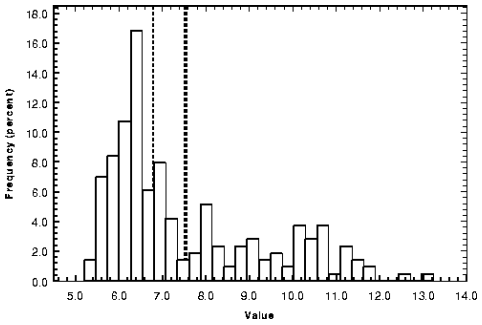
Kriging: *human burden to model numerous variograms*
Simulation: *machine burden to generate many realizations*

Smart Sampling

Example Data Set



214 samples



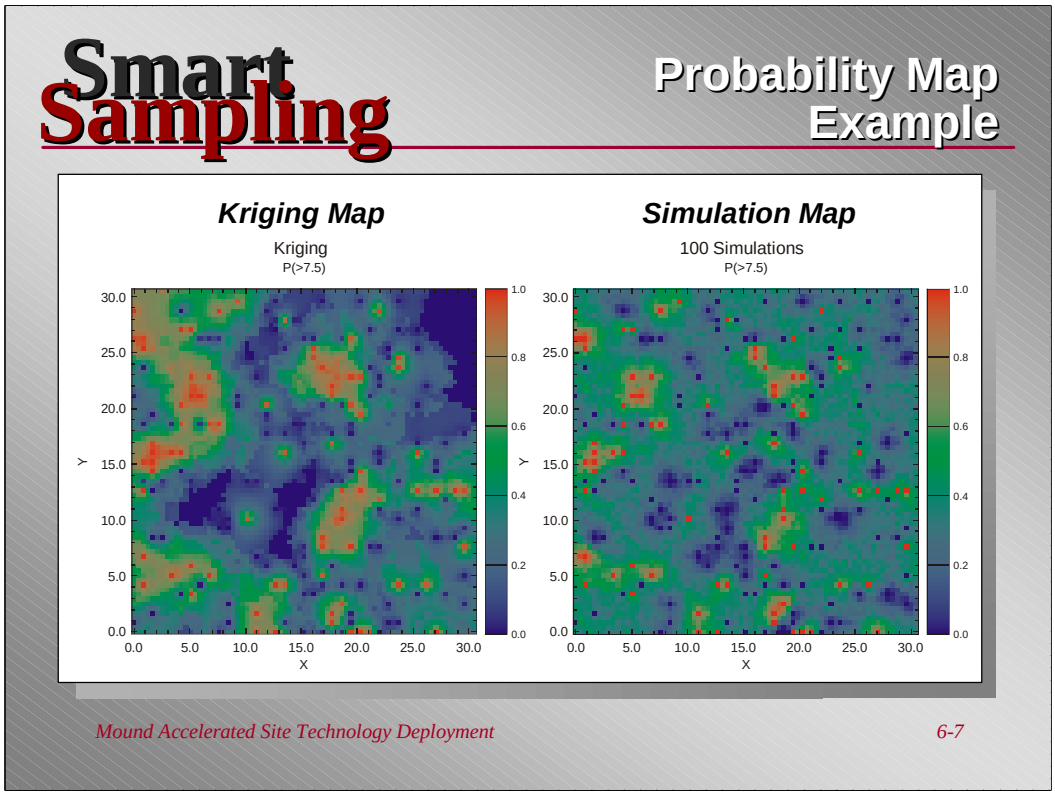
Histogram of data set

Mound Accelerated Site Technology Deployment *6-6*

Data set of 214 permeability samples from Flow Visualization Lab at Sandia.

The mean of the distribution is about 7.5 pCi/g, the median is a little less ~ 6.8. The higher values on the tail affect the mean more than the median.

We will use this data set to create probability maps through both indicator kriging and simulation at a threshold value of 7.5 (UNITS?).

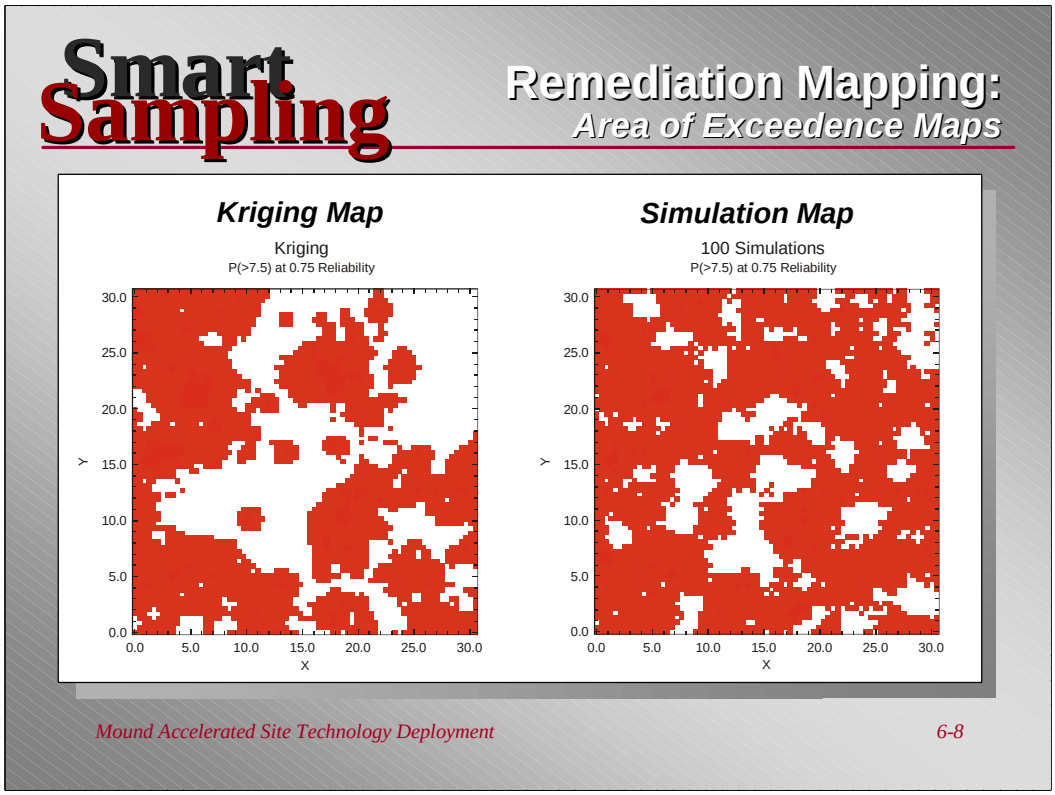


Here we are mapping the probability at every location of exceeding 7.5 pCi/g. Probabilities range from 0 to 1. In some cases, there is a very small probability (blue/green); in some cases, it's very high (red/orange).

The simulation in this case tends to smear out the probabilities, we don't see the distinct zones of 0 that are produced with kriging. If we ran enough of these simulations, the resulting map should look exactly like the kriging map, but since the variograms are somewhat different for the two techniques it doesn't quite happen.

Kriging tends to connect the high and low values considerably better than the simulation. When we first looked at indicator kriging as a spatial estimation technique (Section 4), we assigned values of 0 (below) and 1 (above) in relation to a designated threshold. Now what we're interested in is not simply if the value is above or below, but how close the value is to the threshold, so we apply an indicator transform and then use kriging to get the probabilities (*probability indicator kriging* PIK).

As we go through the process, we will compare how these differing techniques affect our remediation maps and the locations we might choose to take more samples.



The decision rule in this example is to clean up everything ≥ 7.5 pCi/g to 75% reliability. This means that when remediation is complete, up to 25% of the verification samples can come back over threshold and clean up will still be a success.

To generate these remediation maps, we contoured the probability maps at the .75 cut off. The red areas here mark the locations that have a .75 probability of exceeding the action level and must be cleaned up.

In this example, simulation provides a much more conservative estimate of what needs to be cleaned up.

Sean would like to do class exercises using same data set and both techniques, look at differences.

Smart Sampling

Using Probability Maps

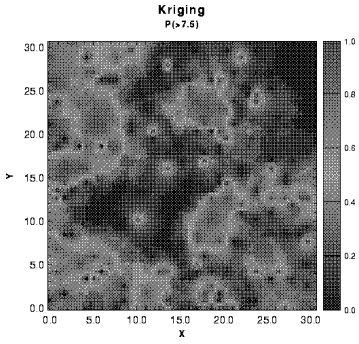
Use probability maps for:

- ***Predictive modeling of the occurrence of the variable (action level)***
- ***Optimizing the location of future sampling.***
- ***Providing the p(fail) term in an economic objective function***

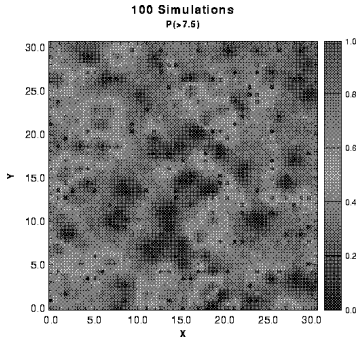
Mound Accelerated Site Technology Deployment *6-9*

Additional samples should be located in the areas of highest uncertainty, the zones at around 0.5 probability of exceedence (green zone).

Kriging Map

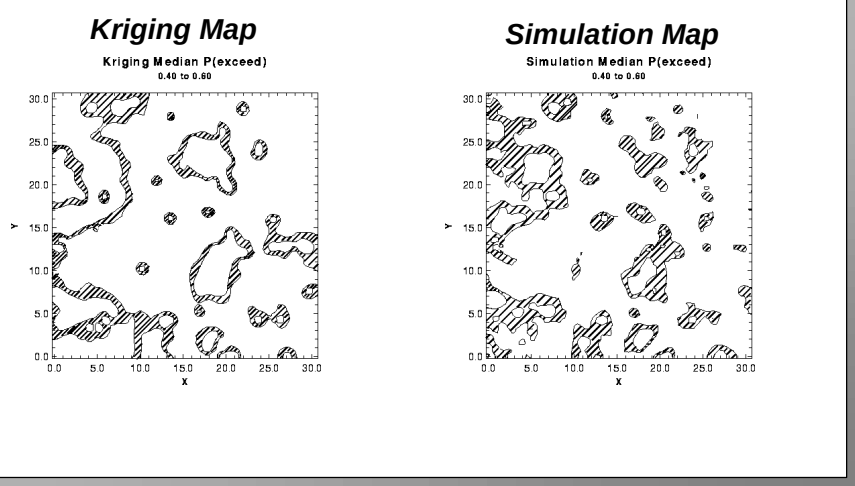


Simulation Map



Smart Sampling

Median Probability as Sample Target



Mound Accelerated Site Technology Deployment *6-10*

In these examples, everything between .4 and .6 probability of exceedence has been contoured and shaded. These are the areas of highest uncertainty, where additional sampling will yield the most information.