# 2-Day Course – Spatial Modeling with Geostatistics

**Prof. Michael J. Pyrcz, Ph.D., P.Eng.**
**Associate Professor**

**Hildebrand Department of Petroleum & Geosystems Engineering**
**University of Texas at Austin**

**Bureau of Economic Geology, Jackson School of Geosciences**
**University of Texas at Austin**

**"In two days, what a geoscientists needs to know about geostatistics, and workflows to get you started with applying geostatistics to impact your work."**

# Spatial Modeling with Geostatistics

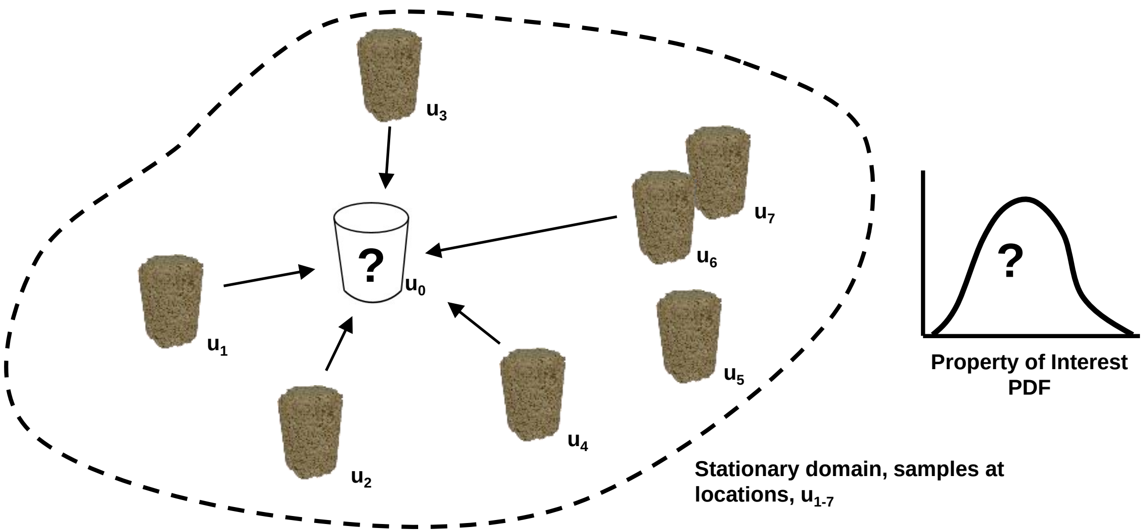
# Uncertainty Management

**Lecture outline . . .**

- **What is Uncertainty?**
- **Calculating Uncertainty**
- **Types of Uncertainty**
- **Uncertainty Workflows**

# How is Uncertainty Represented?

- Uncertainty is not an intrinsic property of the subsurface.
- At every location ($\mathbf{u}_\alpha$) within the volume of interest the true properties  could be measured if we had access (facies, porosity etc.).
- Uncertainty is a function of our ignorance
- This sparsity of sample data combined with heterogeneity results in uncertainty.



# How is Uncertainty Represented?

- Subsurface uncertainty is a model.
- We should use the term “uncertainty model”.
- As with all models, our uncertainty model is imperfect, but useful.
- Our uncertainty assessment is a function of a set of subjective decisions and parameter choices.
- The degree of objectivity is improved by ensuring each of the decisions and parameters are defendable given the available data and judicious use of analogs.

# How is Uncertainty Represented?

**Using Multiple Models:**

- **We represent uncertainty with multiple models.**
- **It is convenient to assume that each model is equiprobable, but one could assign variable probability based on available local information and analogs.**
- **When the input decisions and parameters are changed then these are known as scenarios**
- **When the input decisions and parameters are held constant and only the random number seed is changed then these are known as realizations**

# How is Uncertainty Represented?

- **It is generally not appropriate to analyze a single or few scenarios and realizations.**
- **For decision making, use all the models all the time applied to the transfer function (e.g. volumetric calculation, contaminant transport, ore grade scale up, flow simulation etc.)(Deutsch, “All the Realizations All the Time, 2018)**

# How is Uncertainty Represented?

- **Uncertainty exists because of incomplete data:**
  - **Cannot be avoided**
  - **Can be reduced by consideration of all relevant data, and**
  - **Can be managed / modelled / utilized**
  - **More data may increase uncertainty!**

- **The main steps we will propose are:**
  - **Conceptualize a model for the process**
  - **Quantify the uncertainty in each of the inputs**
  - **Transfer the input uncertainty through to output uncertainty**
  - **Make optimal decision in presence of uncertainty**

# How is Uncertainty Represented?

**Consider Impact on Uncertainty of:**

## 1. Trend Modeling

– **Variance is partitioned between trend, deterministic, known and residual, stochastic, unknown.**

**Variance Components**

$$\sigma^2 = \sigma_t^2 + \sigma_r^2 + \cancelto{0}{2\, C_{t,r}(0)}$$

## 2. Variogram Modeling

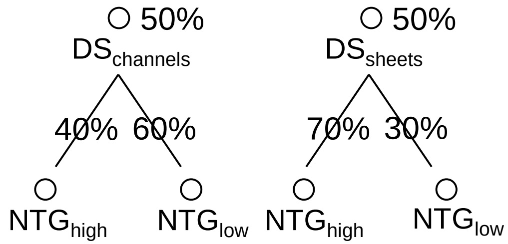
– **Variance is partitioned spatially with variogram features**

**Nested structures each describe spatial uncertainty components.**

$$\sigma_k^2 = \sigma^2_{nugget} + \sigma^2_{spherical} + \sigma^2_{trend}$$

## 3. Scenario / Case Modeling

– **At times it is NOT possible to represent uncertainty as a continuous distribution.**

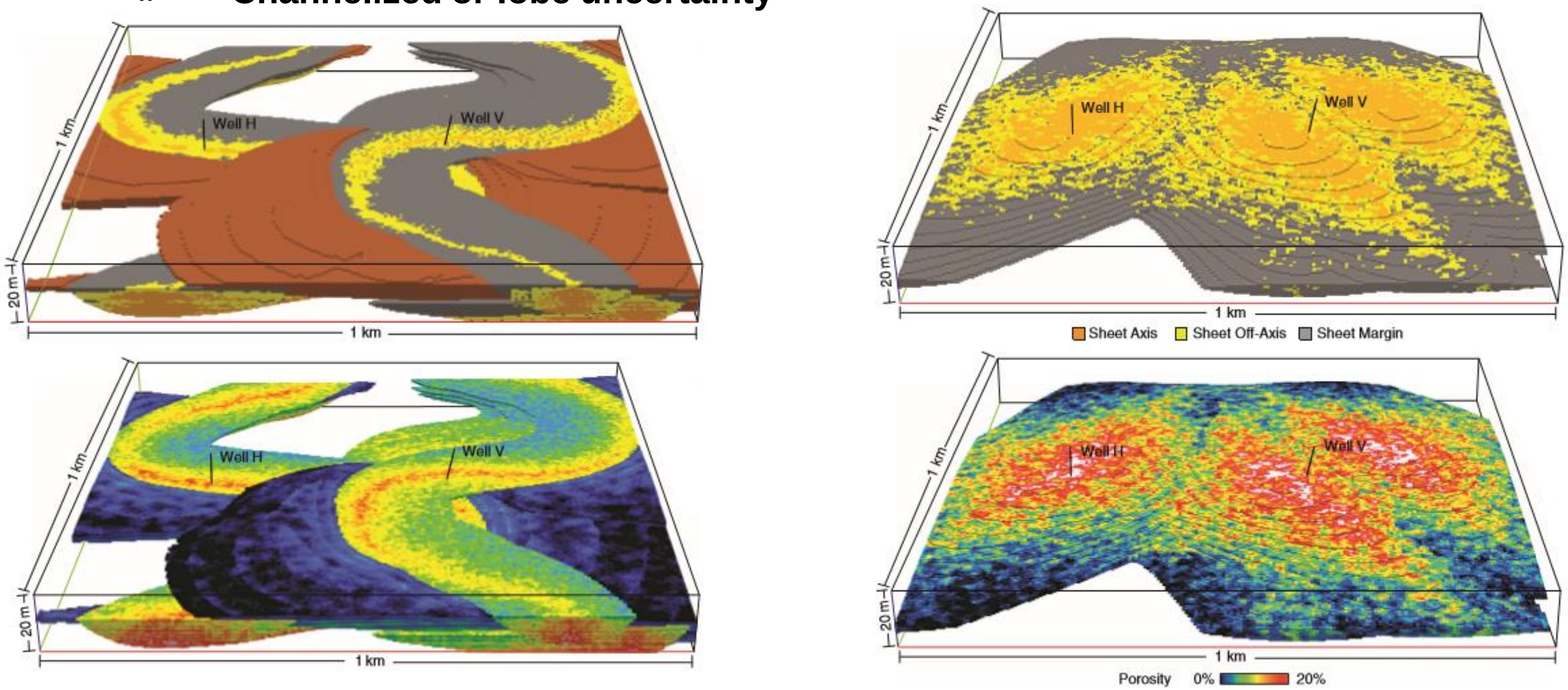



**Assign nested case probabilities.**

**P(NTG,DS) = P(NTG | DS) x P(DS)**

# How is Uncertainty Represented?

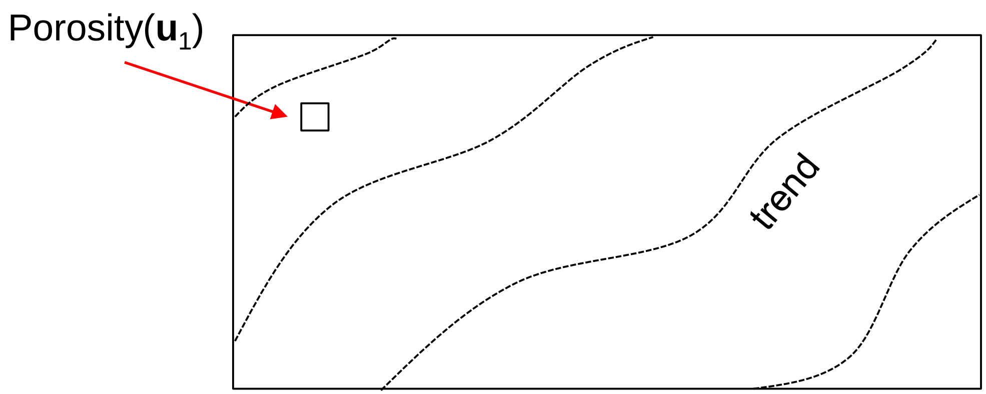
- **Discrete Scenarios**
  - **At times it is NOT possible to represent uncertainty as a continuous distribution.**
  - **Discrete scenarios are required**
    - » **Porosity compaction trend - yes or no.**
    - » **Channelized or lobe uncertainty**



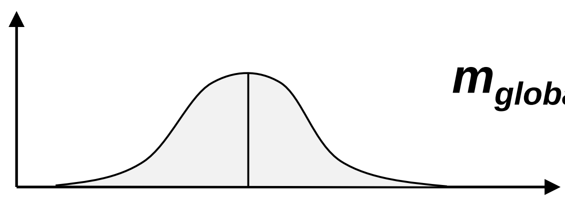

  - **Avoid low, med and high only modeling if possible.**

# How is Uncertainty Calculated?

- **Simple example:**
  - **Uncertainty in porosity estimate at location $\mathbf{u}_1$.**



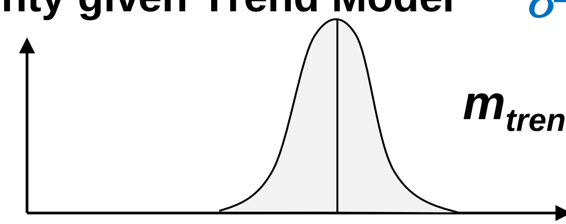

**1. Global Uncertainty from Scenario: Depositional Setting Deepwater Lobe**



$$m_{global}, \sigma^2_{global}\ (\mathrm{Case}_{\mathrm{Lobe}})$$

**2. Local Uncertainty given Trend Model**

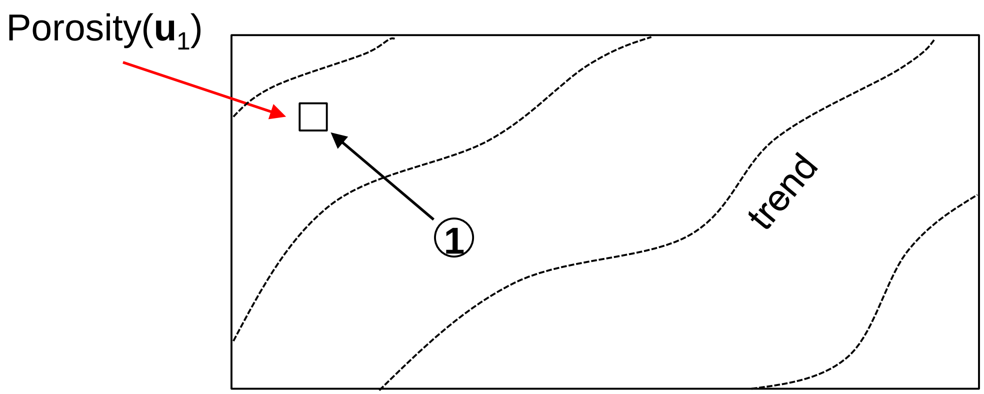
$$\sigma^2_{residual}(\mathbf{u}_1) = \sigma^2_{global}(\mathbf{u}_1) - \sigma^2_{trend}(\mathbf{u}_1)$$



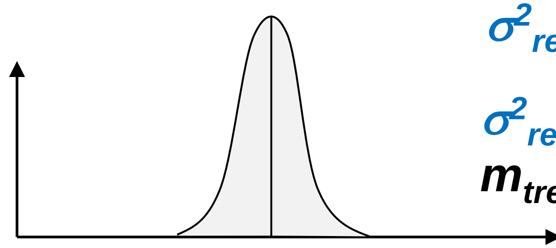
$$m_{trend}(\mathbf{u}_1),\ \sigma^2_{residual}(\mathbf{u}_1)$$

# How is Uncertainty Calculated?

- **Simple example:**
  - **Uncertainty in porosity estimate at location $u_1$.**




**3. Local Uncertainty given Trend Model and Conditioning Data**



$$\sigma^2_{residual\,|\,data}(u_1) \sim \sigma^2_{residual}(u_1) - C(h)$$

$$\sigma^2_{residual\,|\,data}(u_1) = \sigma_k^2$$

$$m_{trend\,|\,data}(u_1) = Z^*$$

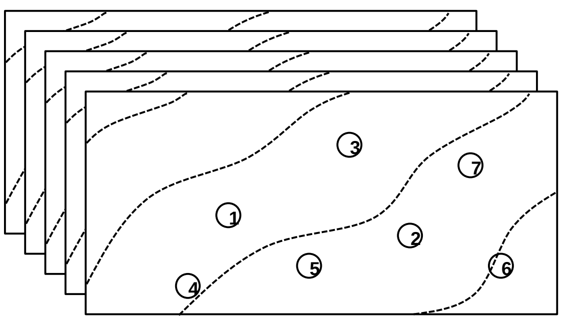
$$m_{trend\,|\,data}(u_1),\ \sigma^2_{residual\,|\,data}(u_1)$$

**Uncertainty at $u_1$ for porosity given** Case$_{Lobe}$**, trend model and conditioning data.**

# How is Uncertainty Calculated?

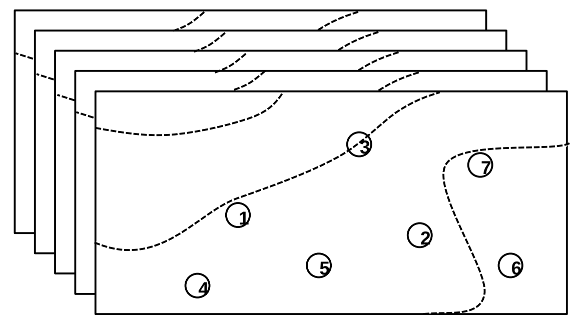
- We have sampled through modeling scenarios and realizations:
  - We can ask any question of the model by considering all scenarios and realizations jointly.

Scenario 1



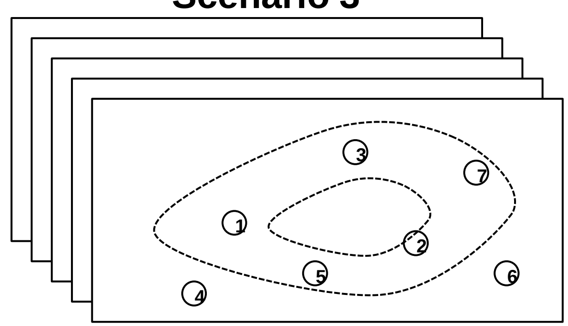

1,…,L realizations (all inputs the same)

Scenario 2




1,…,L realizations (all inputs the same)

Scenario 3




1,…,L realizations (all inputs the same)

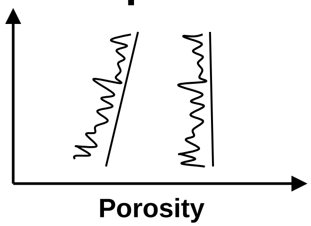
# How is Uncertainty Calculated?

## Modeling Subsurface Uncertainty Without Scenarios?

Michael Pyrcz (@GeostatsGuy, Univ. of Texas at Austin)

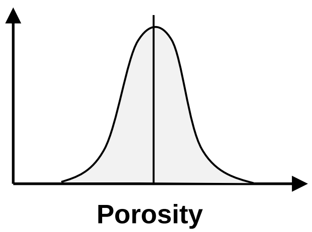
### Implicitly assume perfect knowledge of:

**Primary and Secondary Data**



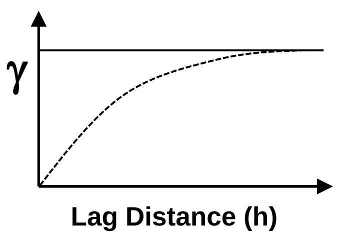

**conditioning data - results in reduced local uncertainty**

**Univariate Distributions**



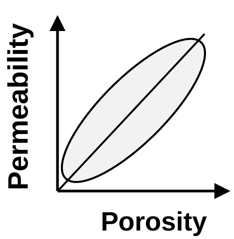

**mean porosity - results in reduced volumetric uncertainty**

**Spatial Continuity**



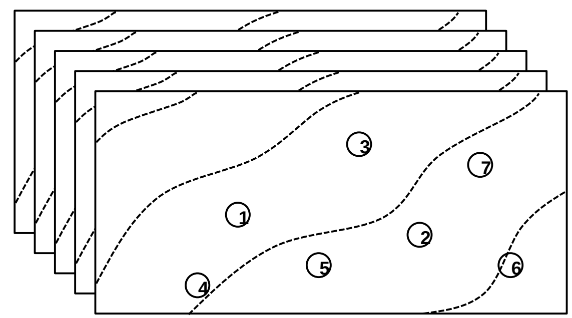

**spatial heterogeneity - results in reduced recovery uncertainty**

**Bivariate Distributions**




**bivariate relationships - results in reduced local uncertainty and recovery uncertainty**

**Scenario 1**



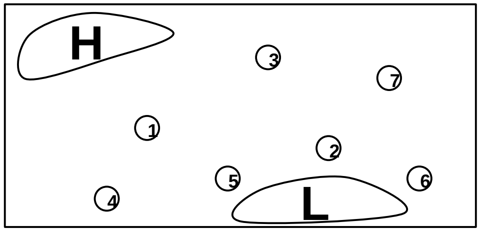

**1,…,L realizations (all the same inputs)**

Relying on ergodic fluctuations to characterize uncertainty in the input statistics.

# How is Uncertainty Calculated?

- **Working without realizations?:**
  - **Implicitly assume perfect knowledge:**

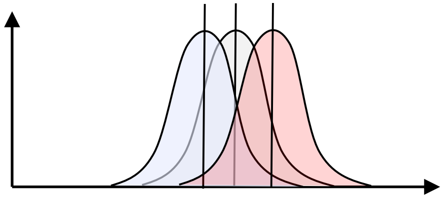
**Spatial Uncertainty**




**Porosity Realization**

**Spatially, away from data. Freeze stochastic islands away from data!**

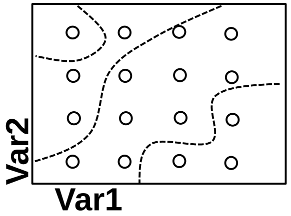
**Input Statistic Fluctuations**



**Porosity Output Distribution**

**Ergodic fluctuations from target statistics.**

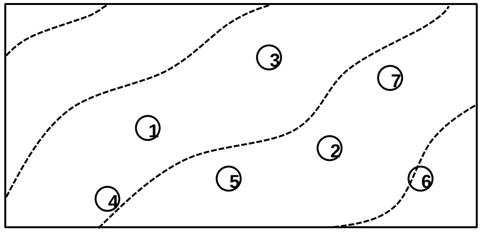
**Response Surface**



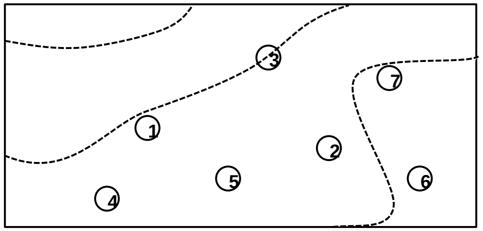

**Response Surface**

**Ergodic fluctuations transfer to response surface creating additional undulations.**

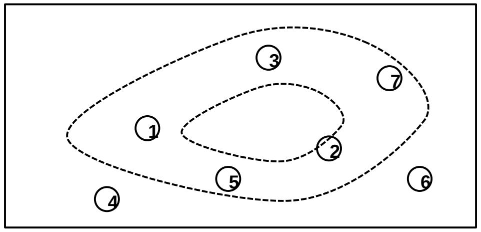
**Scenario 1**



**Scenario 2**

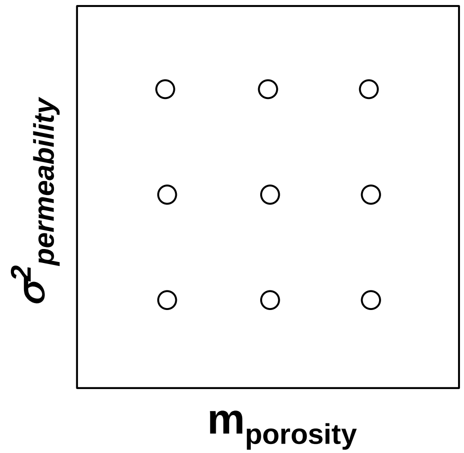


**Scenario 3**



# How is Uncertainty Calculated?

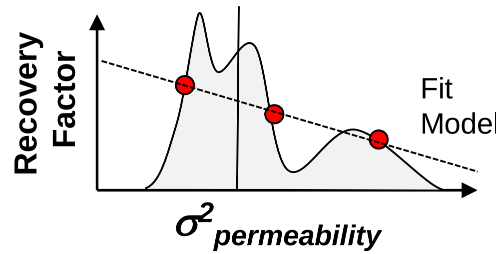
- **Uncertainty Space is Vast!**
  - **Consider typical design**




3 level design, 2 variables
level$^{var}$ = 9 scenarios

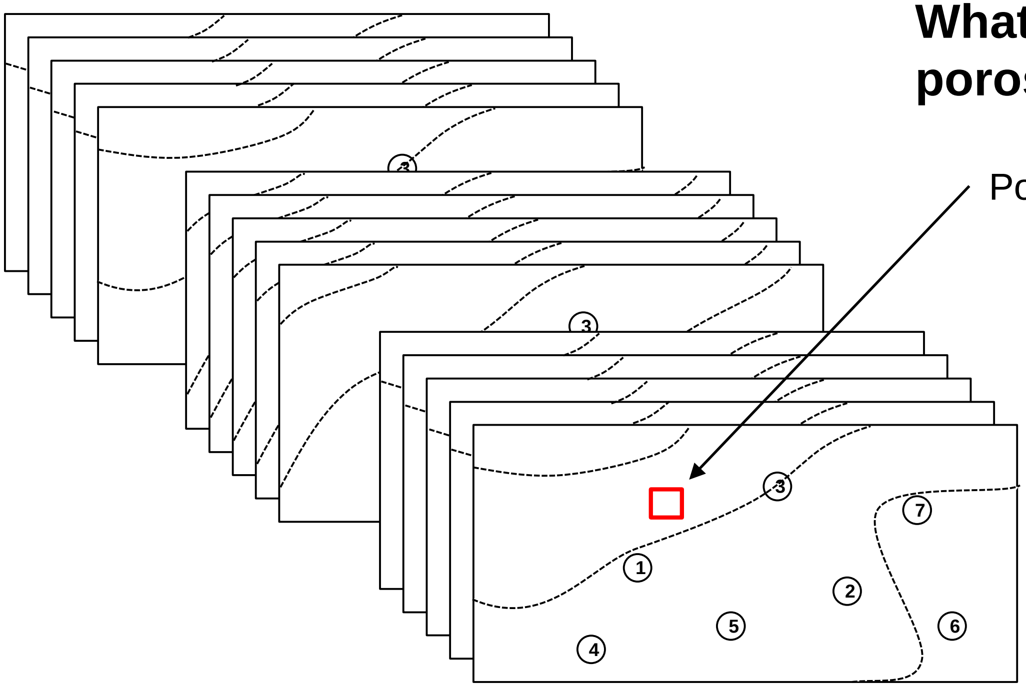
  - **Consider there are typically are around 10 or more uncertain variables.**
    - » **$3^{10}$ = 59,049 scenarios x 10 realizations of each scenario = 590,490 models**
    - » **Variable screening is important!**
    - » **3 level may still poor sampling**

**Univariate Response**



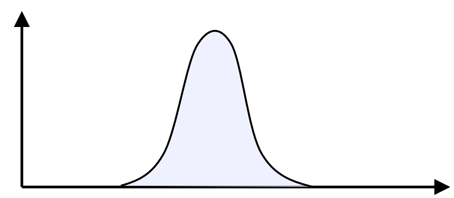
# How is Uncertainty Calculated?

- **We have represented the "uncertainty model" through scenarios and realizations:**
  - **We can ask any question of the model by considering all scenarios and realizations jointly.**




**What is the uncertainty in porosity at location $u_1$?**

Porosity($\mathbf{u}_1$) = F($\Phi$; $\mathbf{u}_1$)

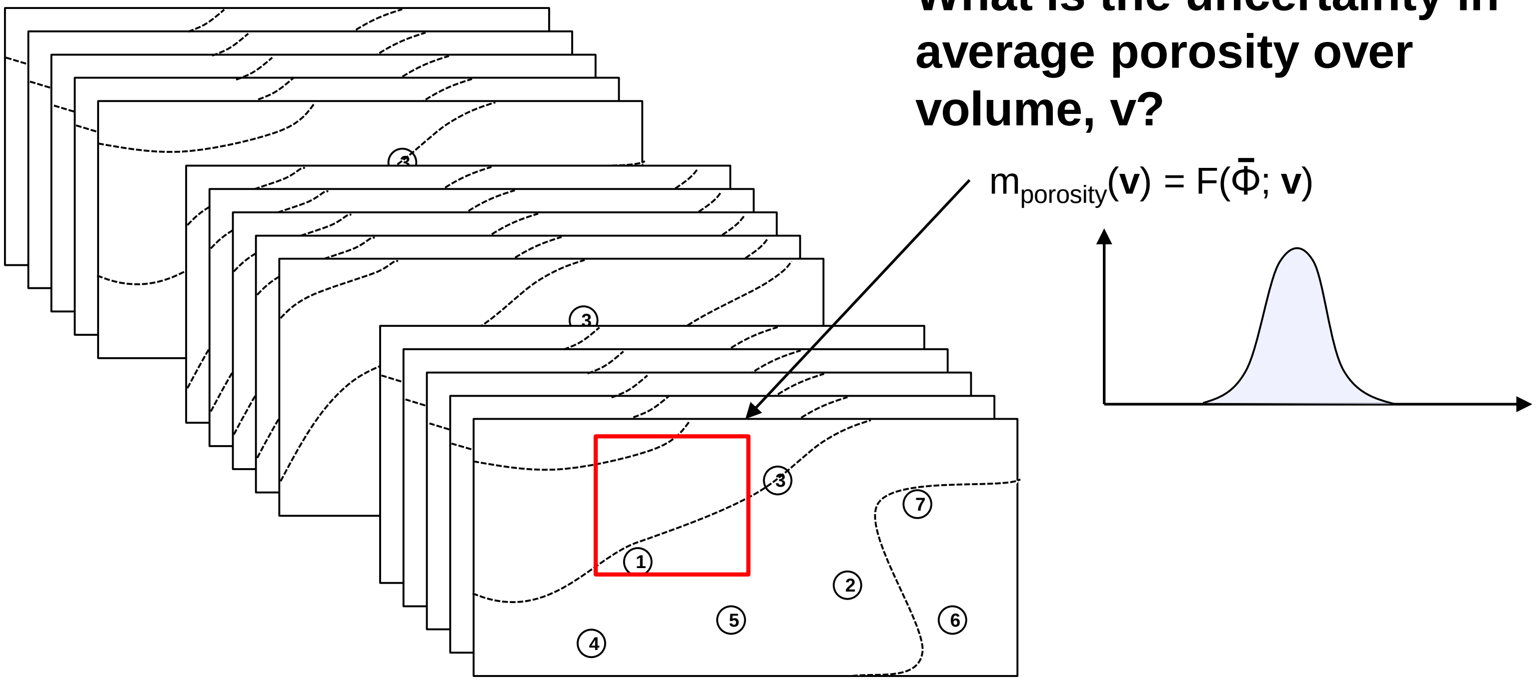


# How is Uncertainty Calculated?

- We have represented the "uncertainty model" through scenarios and realizations:
  - We can ask any question of the model by considering all scenarios and realizations jointly.

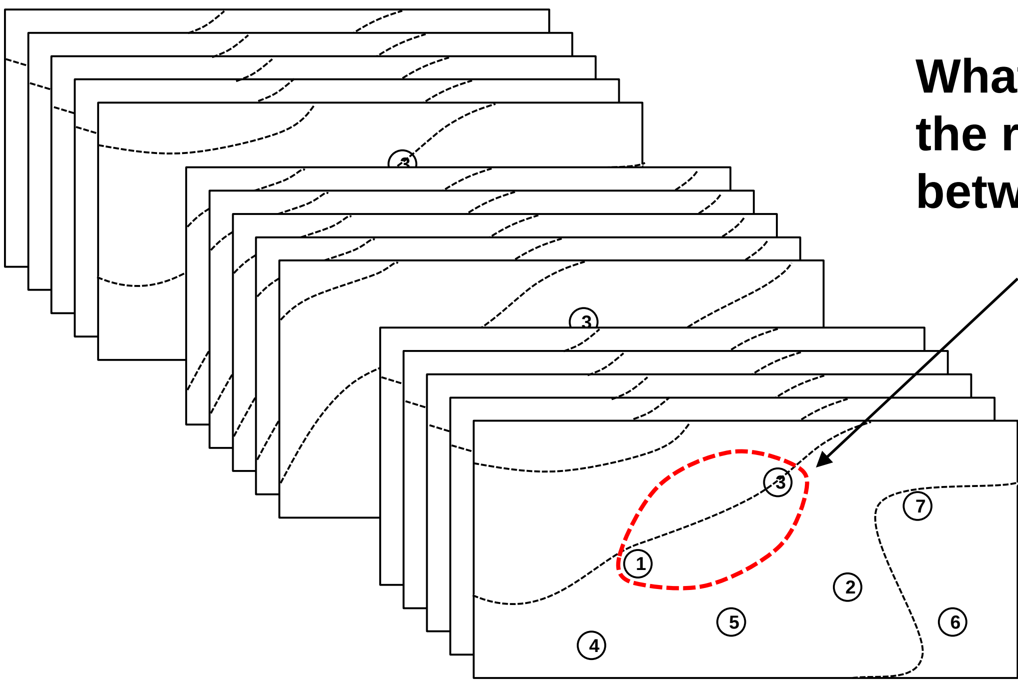
What is the uncertainty in average porosity over volume, v?

$m_{porosity}(\mathbf{v}) = F(\bar{\Phi}; \mathbf{v})$



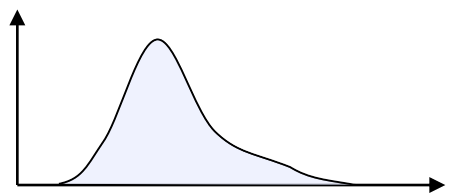
# How is Uncertainty Calculated?

- **We have represented the "uncertainty model" through scenarios and realizations:**
  - **We can ask any question of the model by considering all scenarios and realizations jointly.**



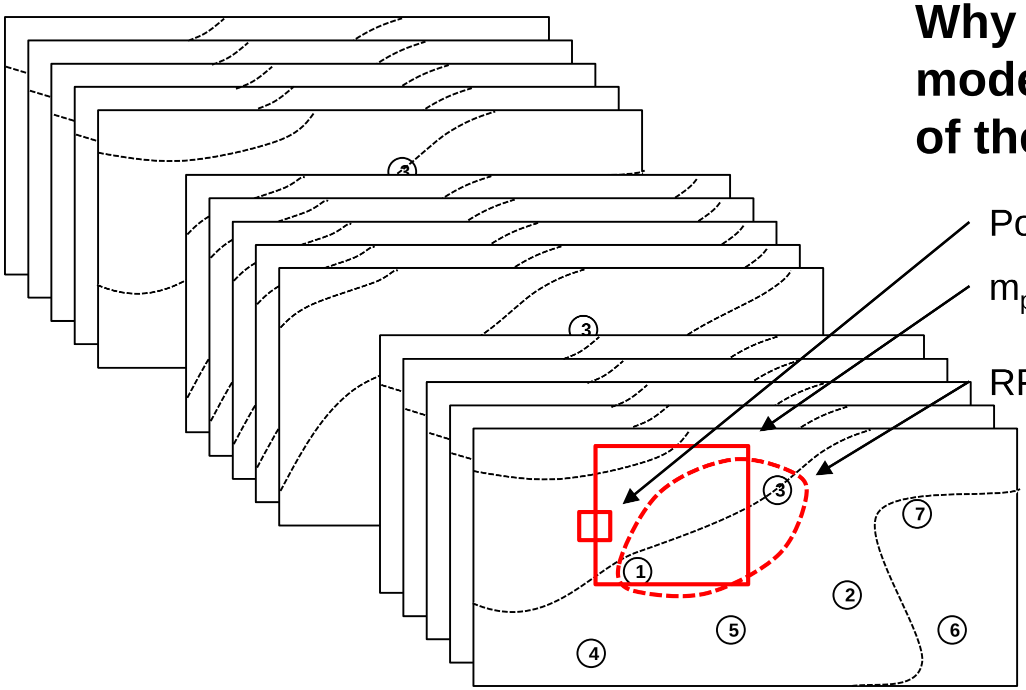

**What is the uncertainty in the recovery factor between $Inj_1$ and $Prod_3$?**

RF(**1-3**) = F(RF; **1-3**)



# How is Uncertainty Calculated?

- **We have represented the "uncertainty model" through scenarios and realizations:**
  - **Models could be ranked as Pxx for a specific question.**
  - **But for every new question, the calculation of rank must be RERUN!**




**Why would a single model be the P90 for all of these questions?**

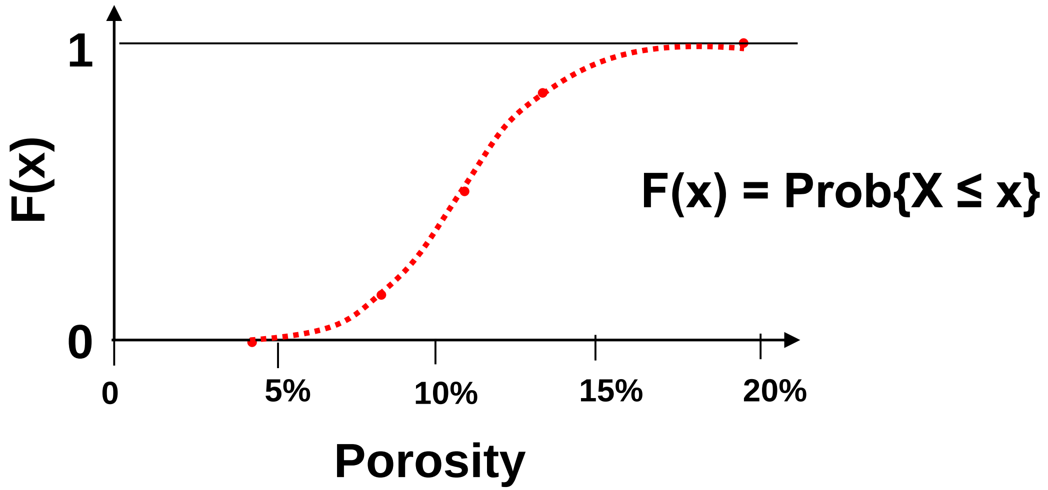
Porosity($\mathbf{u}_1$) = F(Φ; $\mathbf{u}_1$)

$m_{porosity}$($\mathbf{v}$) = F(**Φ**; $\mathbf{v}$)

RF(**1-3**) = F(RF; **1-3**)

# How is Uncertainty Calculated?

- **Frequentist Approach (recall from Lecture 2)**
  - **Sample and pool data samples and formulate a CDF (assumption of stationarity)**
  - **Local data, analog data, calibrated data etc.**




  - **Calculate all required marginal, joint and conditional probabilities density functions and their associated variances or conditional variances**
    - » **$F(x) = \text{Prob}\{X \leq x\}$**
    - » **$F(x,y) = \text{Prob}\{X \leq x, Y \leq y\}$**
    - » **$F(x|y) = \text{Prob}\{X \leq x \mid Y = y\}$**

# How is Uncertainty Calculated?

- **Bayesian Approach (recall from Lecture 2)**
  - **Formulate prior belief**
  - **Calibrate new data, information into a likelihood**
  - **Update to calculate a posterior**

$$\text{Prob}\{A \mid B\} = \frac{\text{Prob}\{B \mid A\} \times \text{Prob}\{A\}}{\text{Prob}\{B\}}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

  - **Induction would be very difficult without Bayesian methods or a lot of data!**
  - E.g. probability of reservoir NTG > 0.5, given < 50 ft. net pay average over wells.

Guidance:

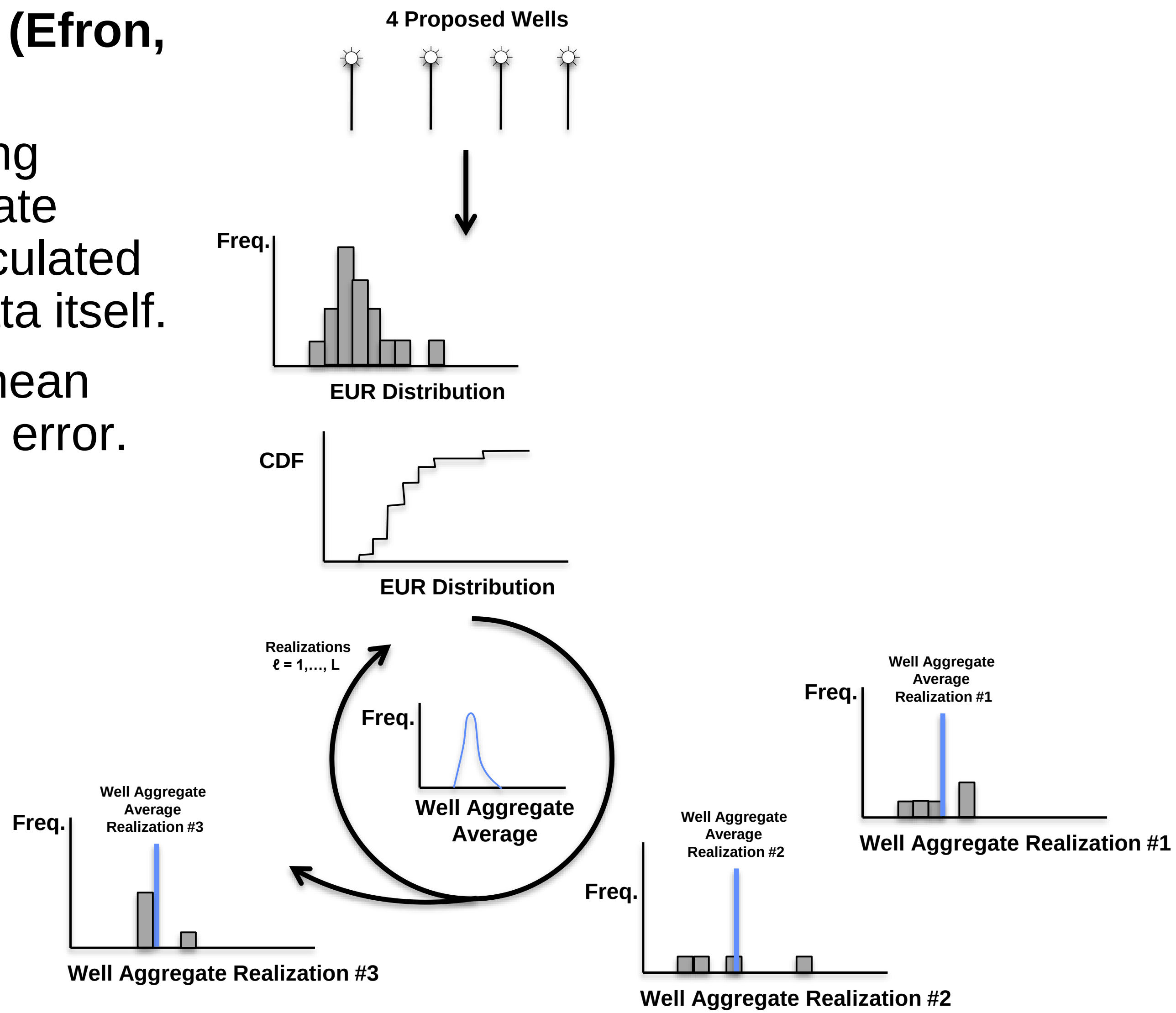
1. Formulate a prior model / marginal probability of event A alone.
   - Prior must be fair, Caers (Stanford) re-iterates that the prior should be quite naïve.
2. Calculate a likelihood that models the impact of event A on B
   - Likelihood must fairly account for this relationship
3. Calculate the evidence term.
   - Often just normalization constant. May be difficult to infer.

# Model Parameter Uncertainty

**Bootstrap Approach (Efron, 1982)**

- Statistical resampling procedure to calculate uncertainty in a calculated statistic from the data itself.
- For uncertainty in mean solution is standard error.

$$\sigma_x^2 = \frac{\sigma_s^2}{n}$$

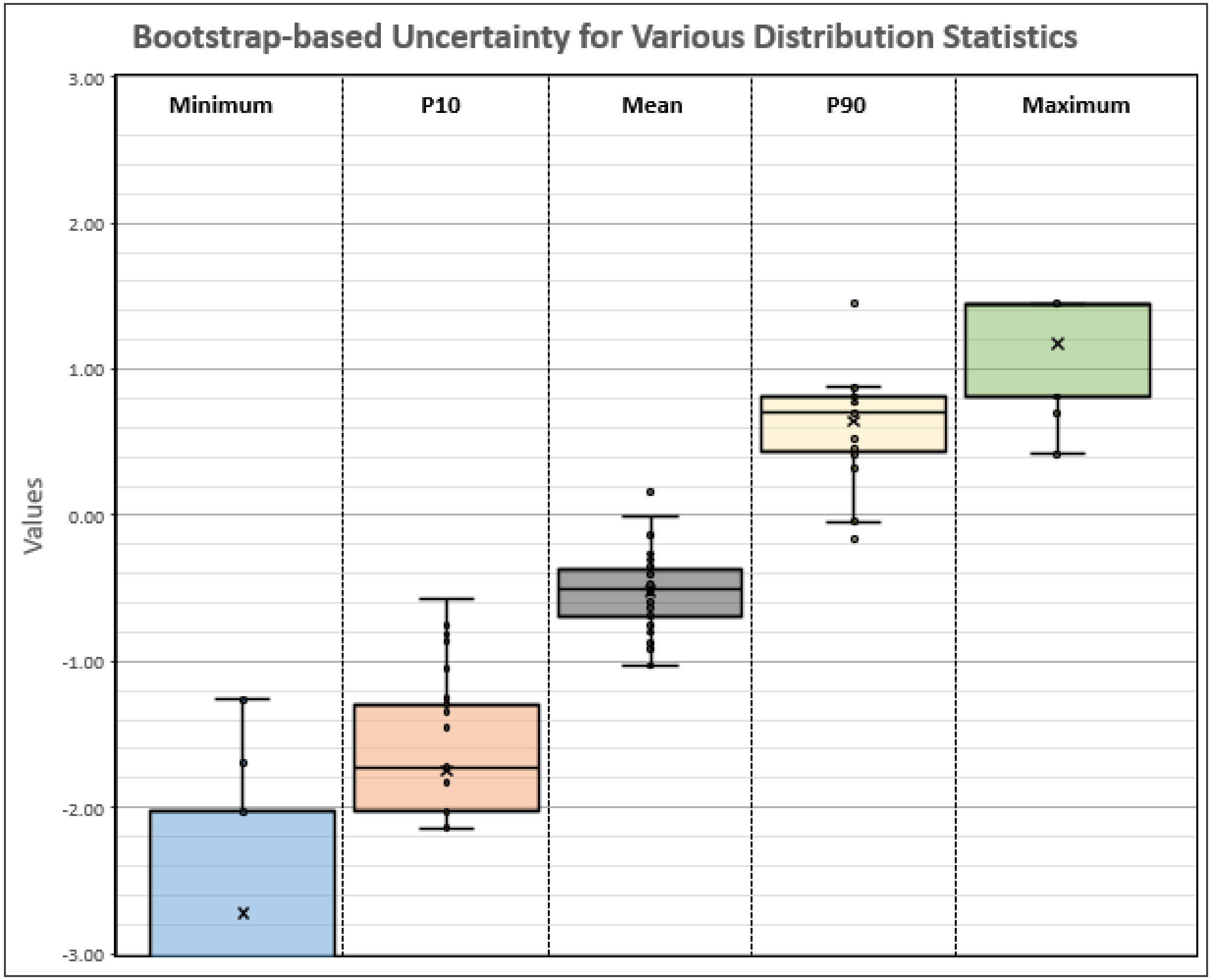


# Bootstrap Hands-on In Excel

**Observed the bootstrap method for estimating uncertainty in a variety of statistics**

What happens if you decrease the number of samples form 20 to 10?

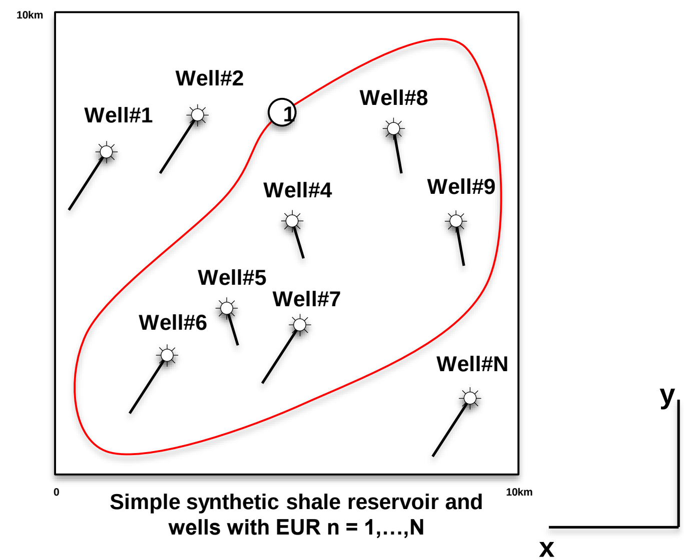
File is Bootstrap_Demo.xlsx

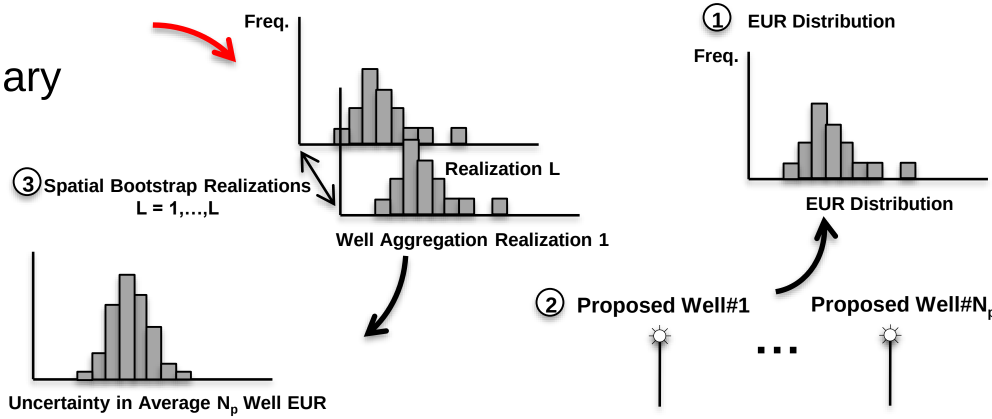


# Model Parameter Uncertainty

**Spatial Bootstrap Approach (Journel, 1994)**

- Spatial bootstrap is a variant of bootstrap that accounts for spatial correlations when resampling
- method accounts for:
  - Proposed and previous well locations
  - Spatial Continuity
  - Local trends and secondary data
- Based on sequential simulation from EUR distribution

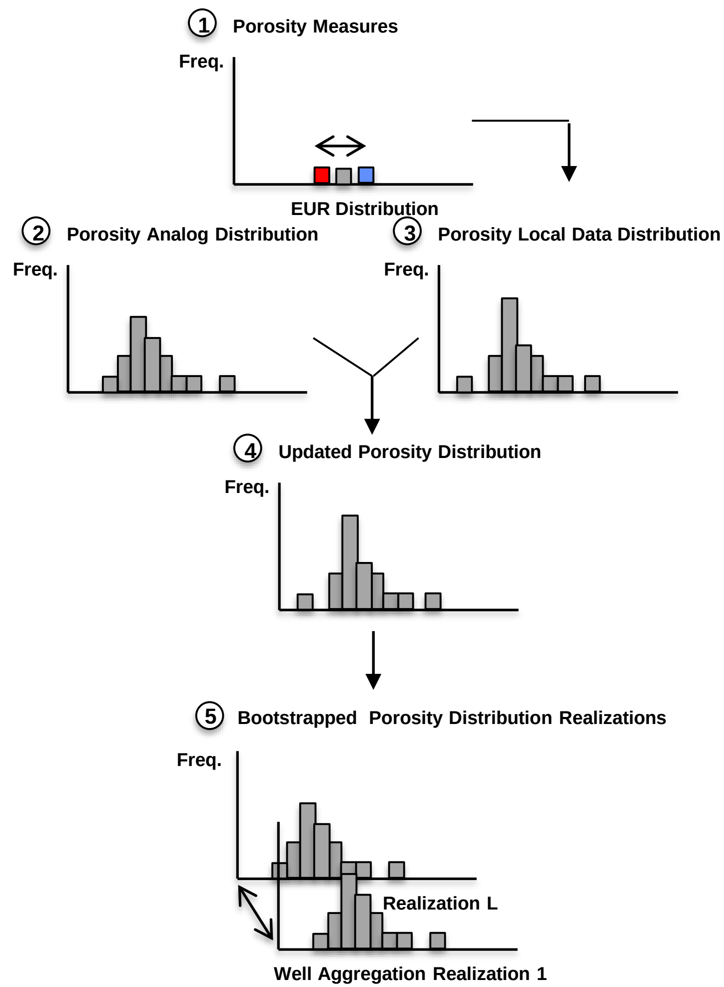


# Model Parameter Uncertainty

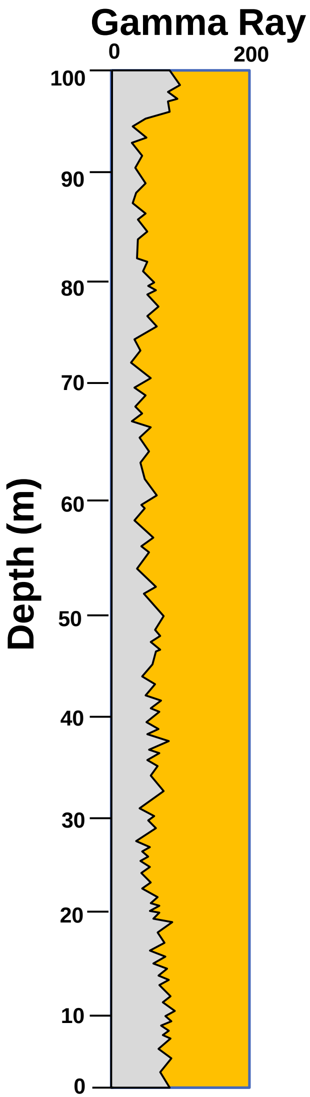
**Example Workflow**

1. Calculate measurement error for all data and calculate data realizations. Insignificant.
2. Frequentist approach to build porosity global CDF from available analogs
3. Frequentist approach to build declustered local data CDF.
4. Bayesian updating of analog prior porosity global CDF
5. Spatial bootstrap to calculate multiple realizations of porosity distribution.



# Types of Uncertainty

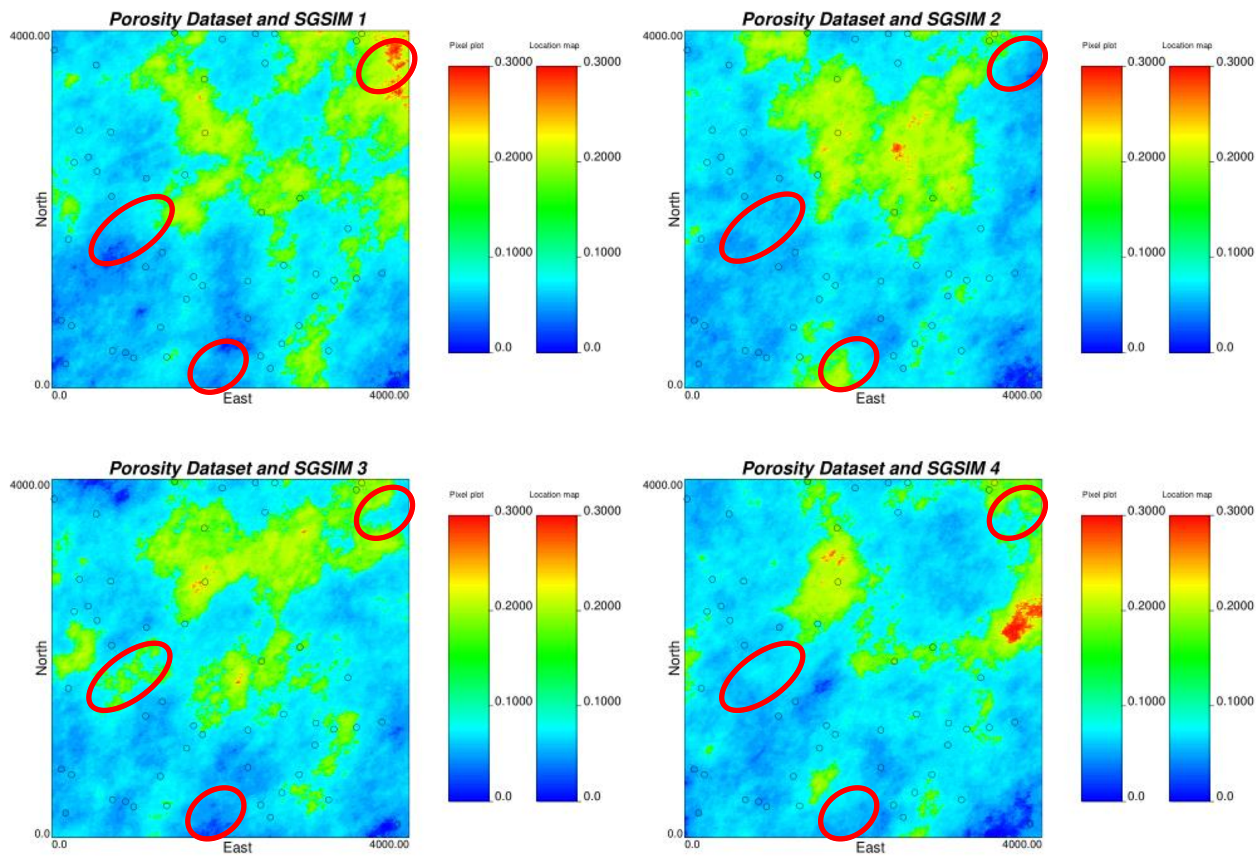
- **Measurement / Interpretation Error.**
  - **Formation evaluation – tool tolerance, calibration error, approximations / assumptions**
  - **Interpreter experience and prior model / assumptions**
  - **How to integrate it?**
    - » **Indicator method code as soft inputs**
    - » **Multiple data realizations in design of experiments**




Example well log.

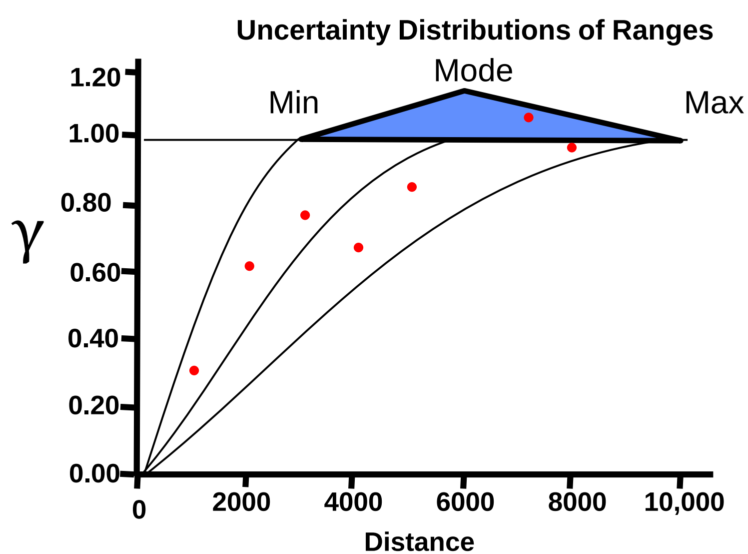
# Types of Uncertainty

- **Spatial Uncertainty**
  - Uncertainty due to spatial offset from sampled locations
  - Integrate through multiple local realizations and scenarios

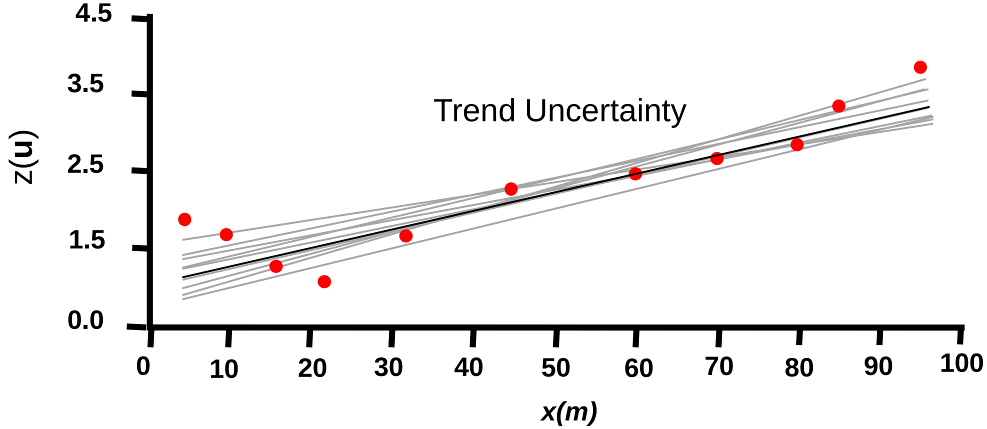


# Types of Uncertainty

- **Parameter Uncertainty**
  - **Uncertainty in the input statistics**
  - **E.g. global reference porosity distribution for simulation**
  - **Formulate distribution scenarios (could bootstrap for parameter realizations)**



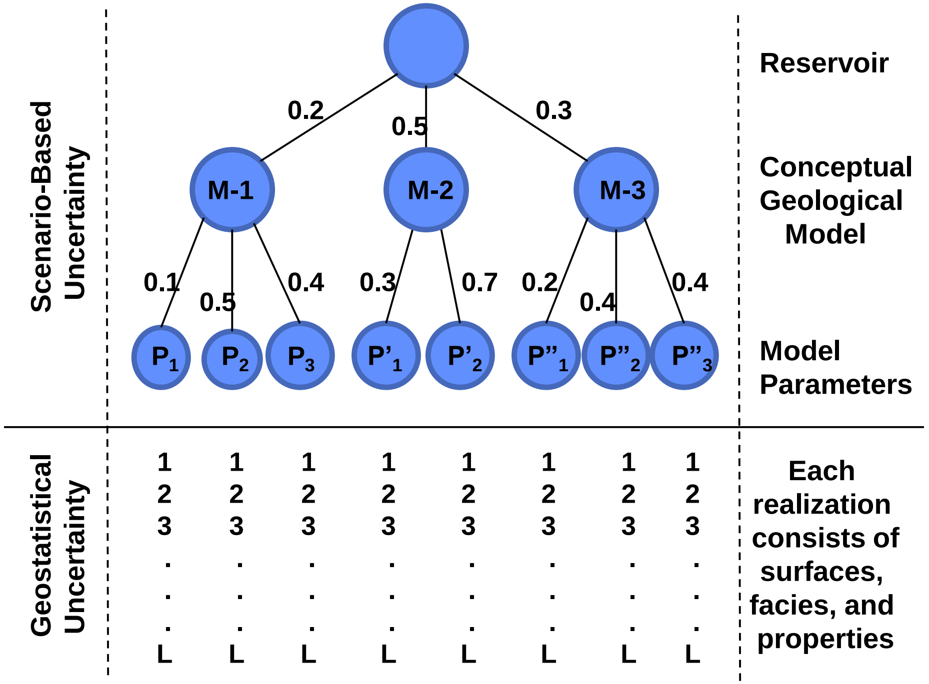

Uncertainty in Porosity Variogram




Trend Uncertainty

# Uncertainty

- **Best Practice**
  1. Seek out all significant uncertainty sources
  2. Assign scenarios when needed with associated de-biased probabilities
  3. Include data realizations if needed.
  4. Also include stochastic realizations to account for spatial uncertainty.
  5. Need enough models the uncertainty space is vast.
  6. Document / defend choices.
  7. Make decision with the entire suite of models



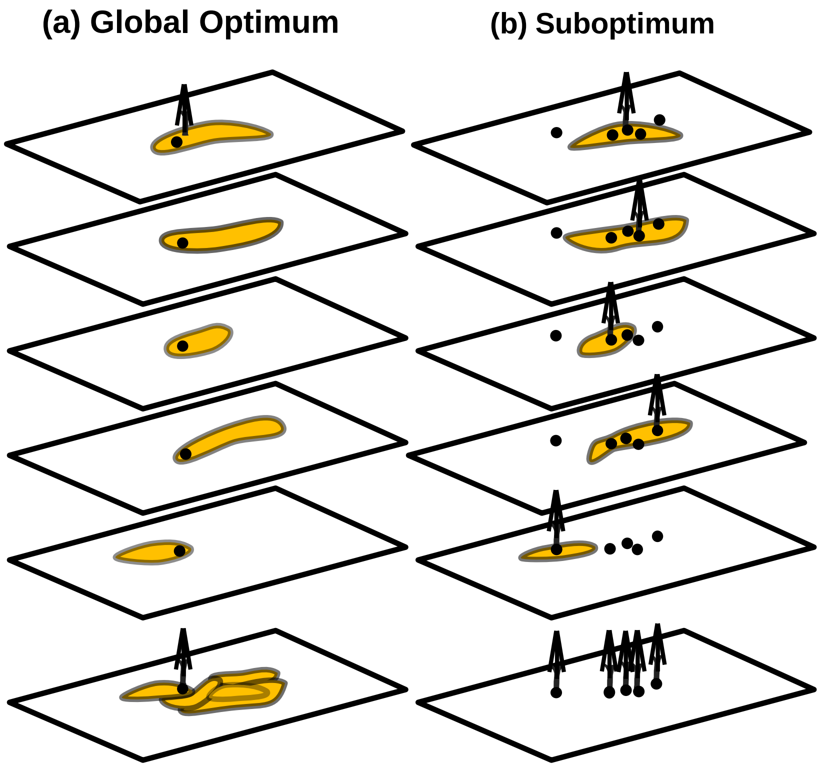

Uncertainty exploration scheme

# Working with Multiple Realizations

- **Making Decisions with Multiple Realizations**

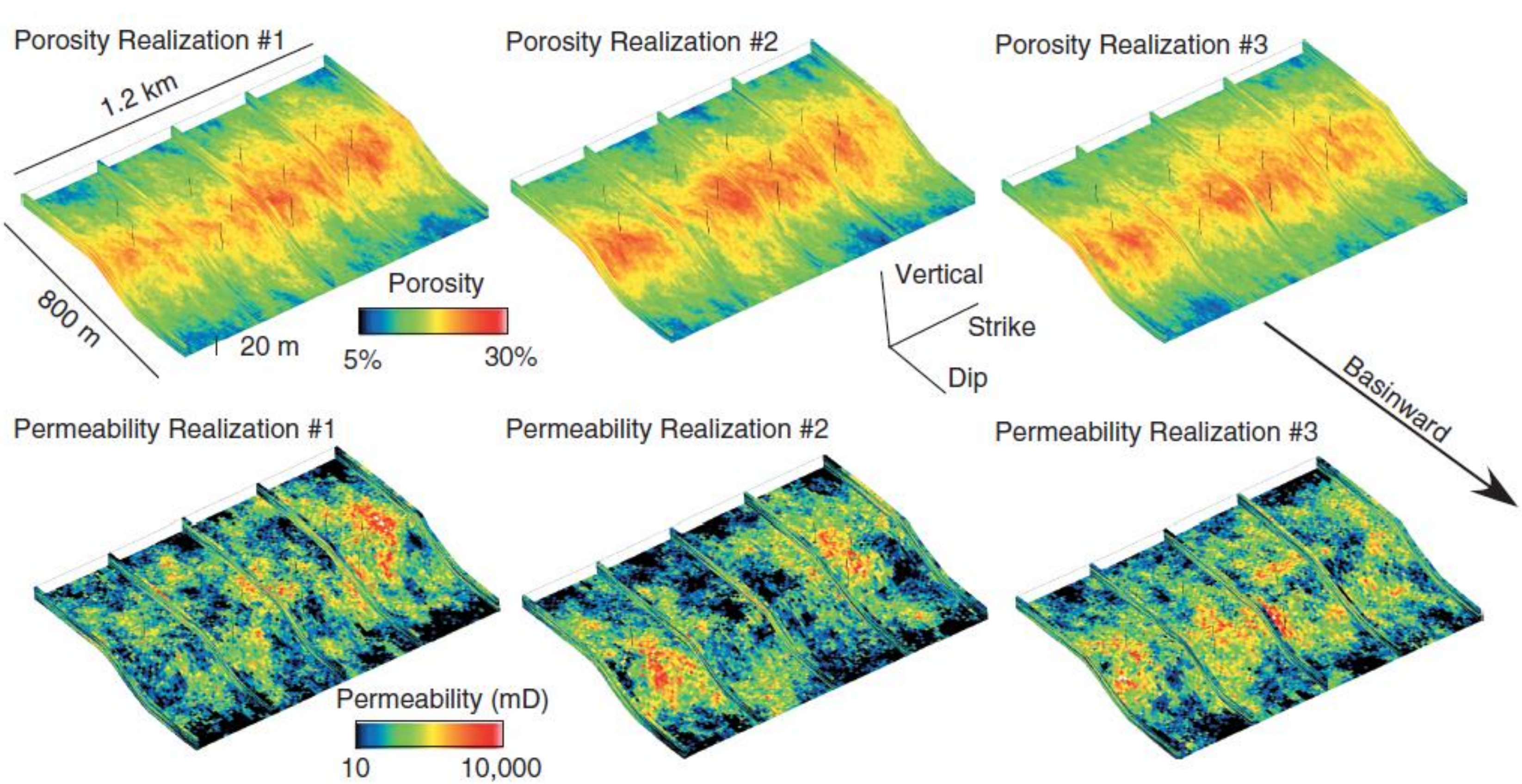
- **One Realizations at a Time**
  - **For each model you could make a suboptimal decision.**
  - **For L models, you would have L suboptimal decisions**

- **All Realizations Jointly**
  - **A single decision taken with all the realizations**
  1. **Place well, for each model calculate the resulting production, and calculate the expected production**
  2. **Repeat for multiple locations**
  3. **Select the location that maximizes production**



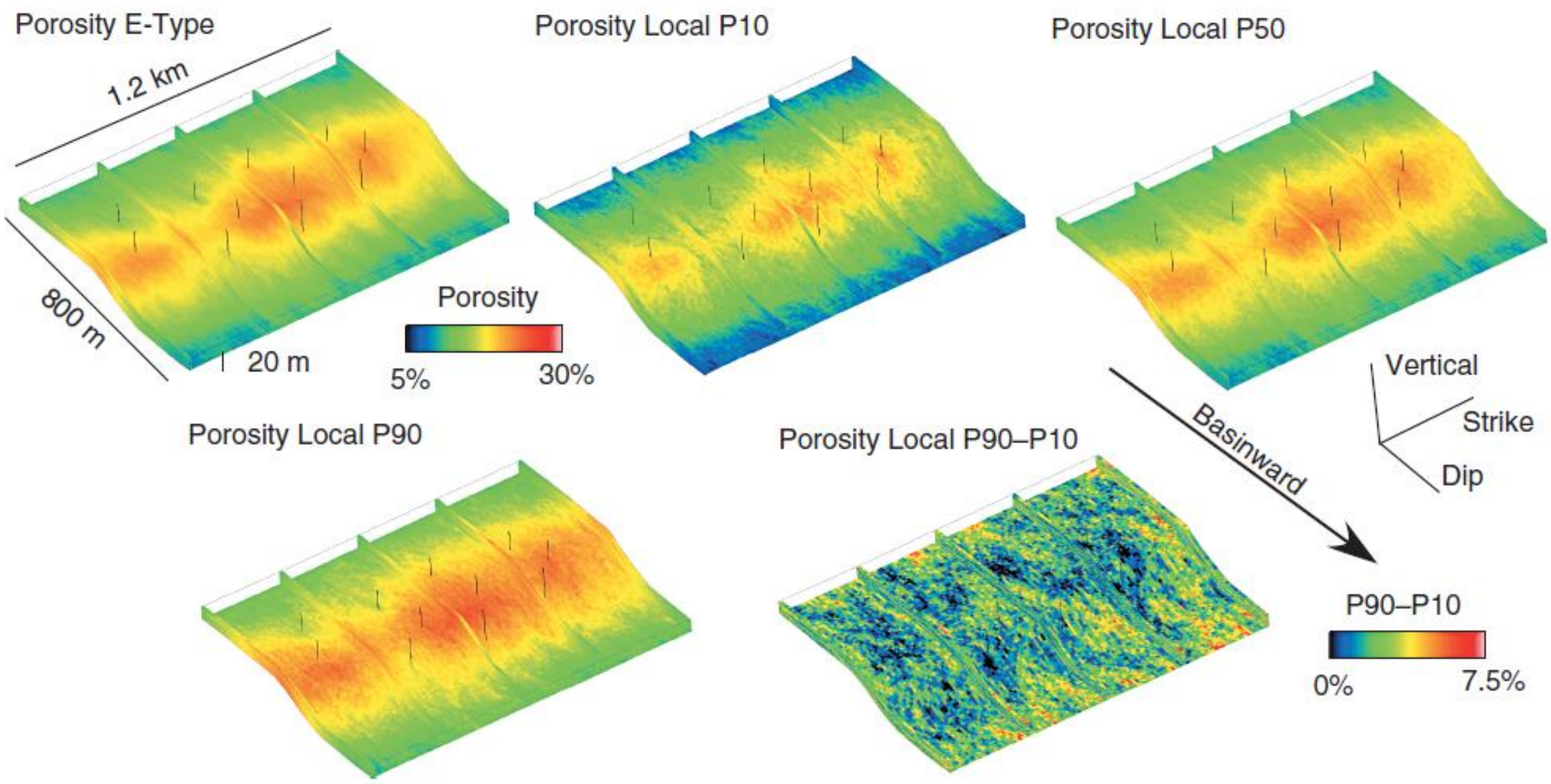
# Working with Multiple Realizations

- **Summarizing over Multiple Realizations**
  - **Multiple reservoir model realizations.**



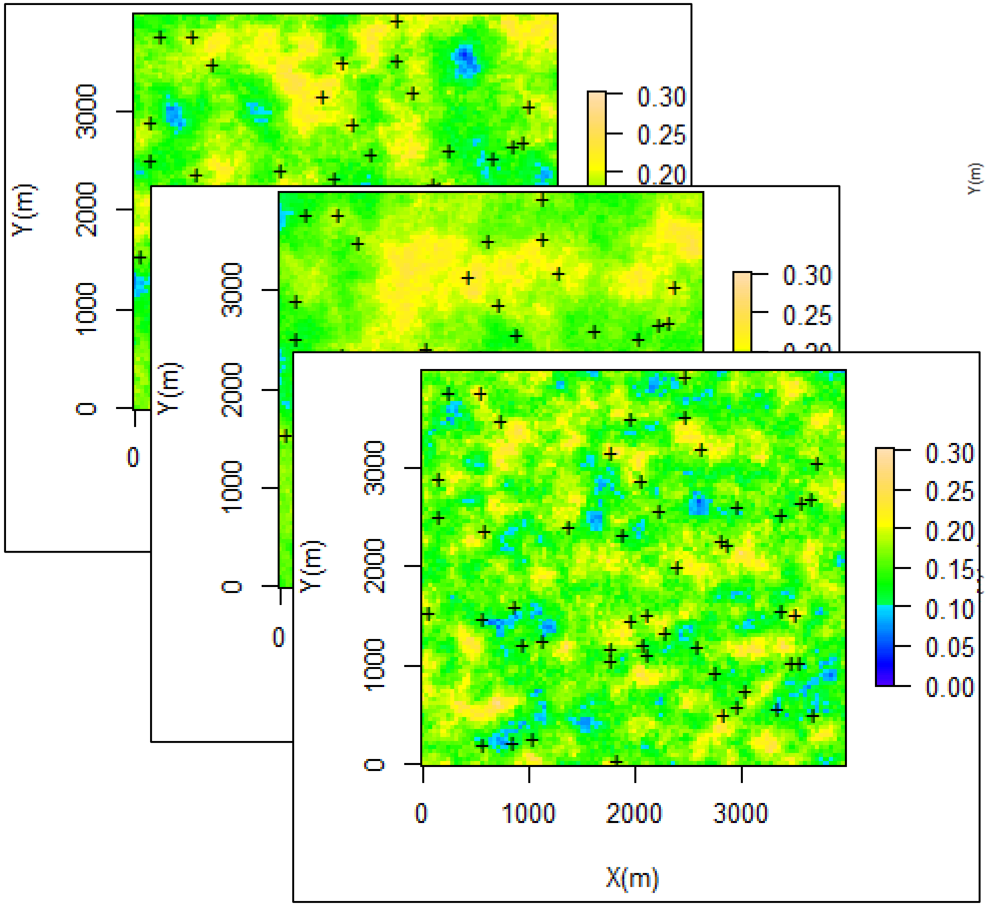
# Working with Multiple Realizations

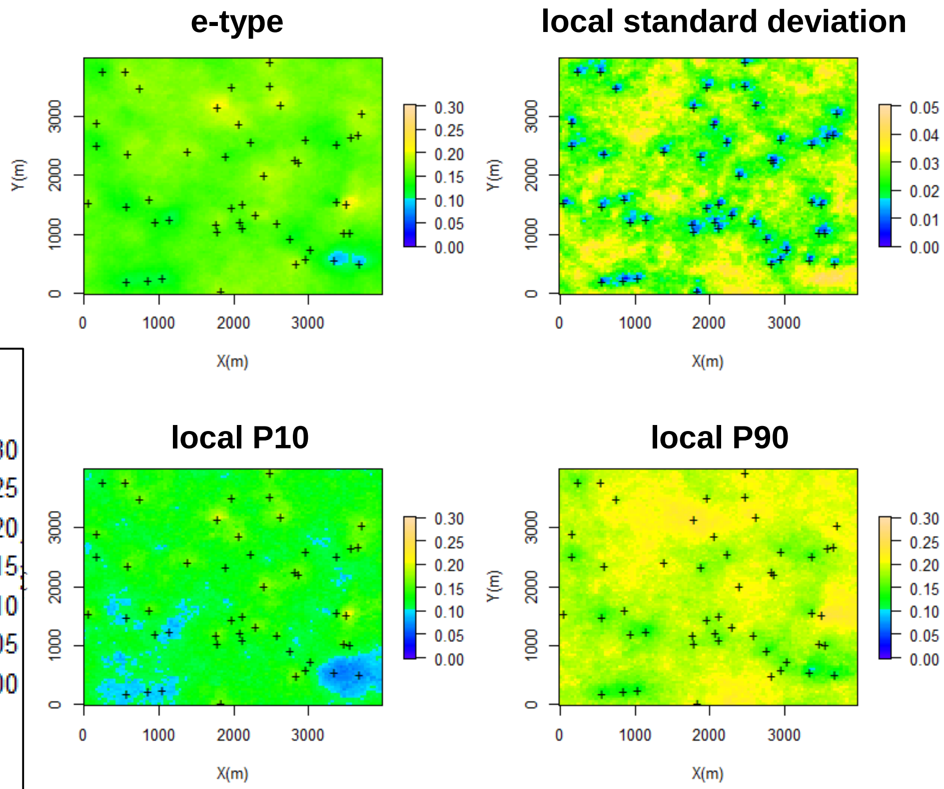
- **Summarizing over Multiple Realizations**
  - **E-type, Local P-value, Local Range, Local Standard Deviation**



# Working with Multiple Realizations Demo in R

- **Summarizing over Multiple Realizations**
  - **3 scenarios, 10 realizations of each**
  - **E-type, Local P-value, Local Range, Local Standard Deviation**



# Working with Multiple Realizations Hands-on in R

- At the very end of the demo, these lines modify the variogram model, rerun the ordinary kriging and plot the estimate and estimation variance.

```
# Three variogram scenarios
vm1 <- vgm(psill = 1.0*sill, "Sph", 600, anis = c(000, 1.0),nugget=0.0*sill)
vm1
vm2 <- vgm(psill = 1.0*sill, "Sph", 1200, anis = c(000, 1.0),nugget=0.0*sill)
vm2
vm3 <- vgm(psill = 0.99*sill, "Sph", 300, anis = c(000, 1.0),nugget=0.01*sill)
vm3
```

Some ideas to explore:

1. Set the nugget high (use nugget parameter and set psill to 1-nugget).

2. Set the range lower and higher (currently 300, 600, 1200) in the function call.

3. Increase and decrease the anisotropy ratio (currently 1.0, isotropic) in the function call.

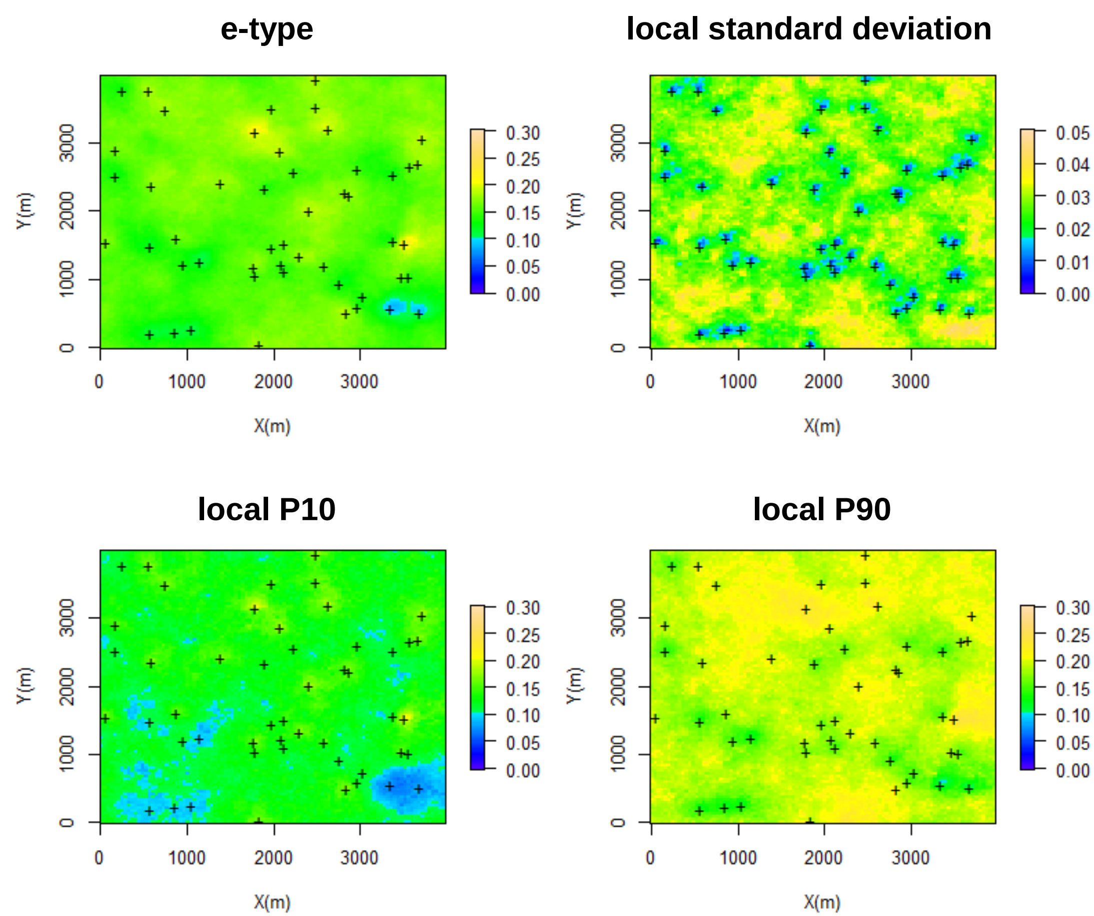
# Working with Multiple Realizations Hands-on in R

- **Where would you drill?**

  1. **To maximize production?**

  2. **To minimize risk of a bad well?**

  3. **To gather information?**



# Summary on Uncertainty

**Calculating Uncertainty in a Modeling Parameter:**

- Use Bayesian methods, spatial bootstrap etc. You must account for the volume of interest, sample data quantity and locations, and spatial continuity.

**If You Know It, Put It In.**

- Use expert geologic knowledge and data to model trends. Any variability captured in a trend model is known and is removed from the unknown, uncertain component of the model. Over fit trend will result in unrealistic certainty.

**Types of Uncertainty:**

- (1) data measurement, calibration uncertainty, (2) decisions and parameters uncertainty, and (3) spatial uncertainty in estimating away from data. Your job is to hunt for and include all significant sources of uncertainty.

# Summary on Uncertainty

## What about Uncertainty in the Uncertainty?

- Don't go there. Use defendable choices in your uncertainty model, be conservative about what you known, document and move on. Matheron taught us to strip all away all defenseless assumptions. Journel warned us to avoid the circular quest of uncertainty in uncertainty in...

## Uncertainty Depends on Scale.

- It is much harder to predict a property of tea spoon vs. a house-sized volume at a location ($u_\alpha$) in the subsurface. Ensure that scale and heterogeneity are integrated.

## You Cannot Hide From It.

- Ignoring uncertainty assumes certainty and is often a very extreme and dangerous assumption.

## Decision Making with Uncertainty.

- Apply all the models to the transfer function to calculate uncertainty in subsurface outcome to support decision making in the presence of uncertainty.

# Uncertainty
# New Tools

| Topic | Application to Subsurface Modeling |
|---|---|
| **Partition Variance / Uncertainty Spatial** | Modeling a trend<br><br>*Use additivity of variance to communicate decomposition of variance into known trend and unknown residual.* |
| **Integrate Scenarios** | Build multiple scenarios<br><br>*Capture the uncertainty in model choices, parameters* |
| **Integrate Realizations** | Build multiple realizations<br><br>*Capture spatial uncertainty away from sampled locations in the subsurface.* |
| **Parameter Uncertainty** | Calculate subsurface parameter uncertainty<br><br>*Utilize bootstrap to assess the uncertainty in an model parameter given sparse sampling.  Consider spatial bootstrap if data is spatially correlated.* |