Uncertainty

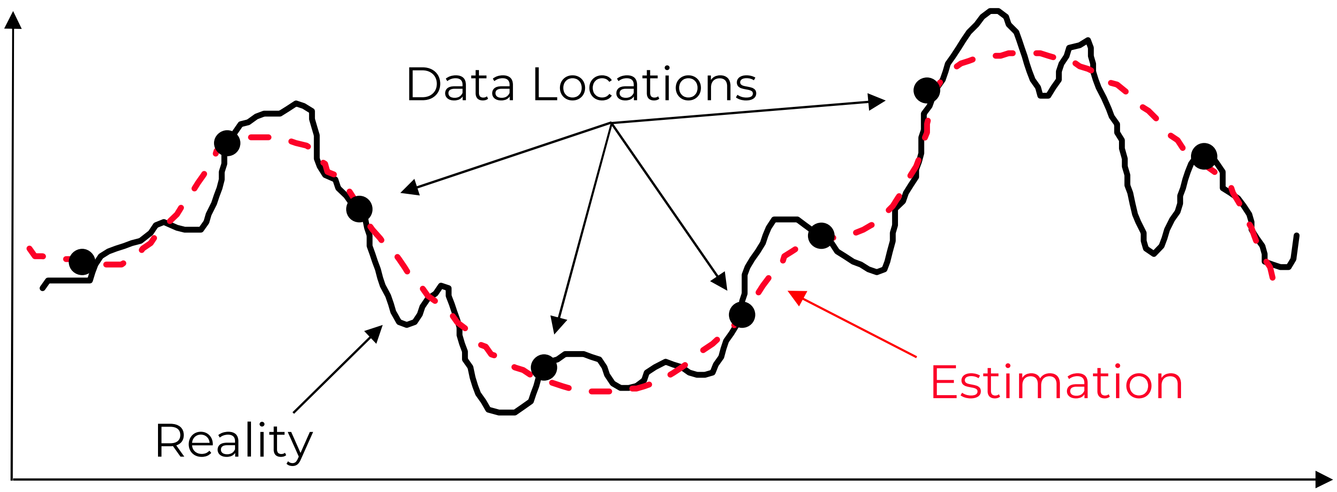
Lecture outline . . .

- Stochastic Simulation
- Sources of Uncertainty
- Representing Uncertainty
- Summarizing Over Realizations
- Uncertainty Summary

STOCHASTIC SIMULATION

MOTIVATION FOR SIMULATION

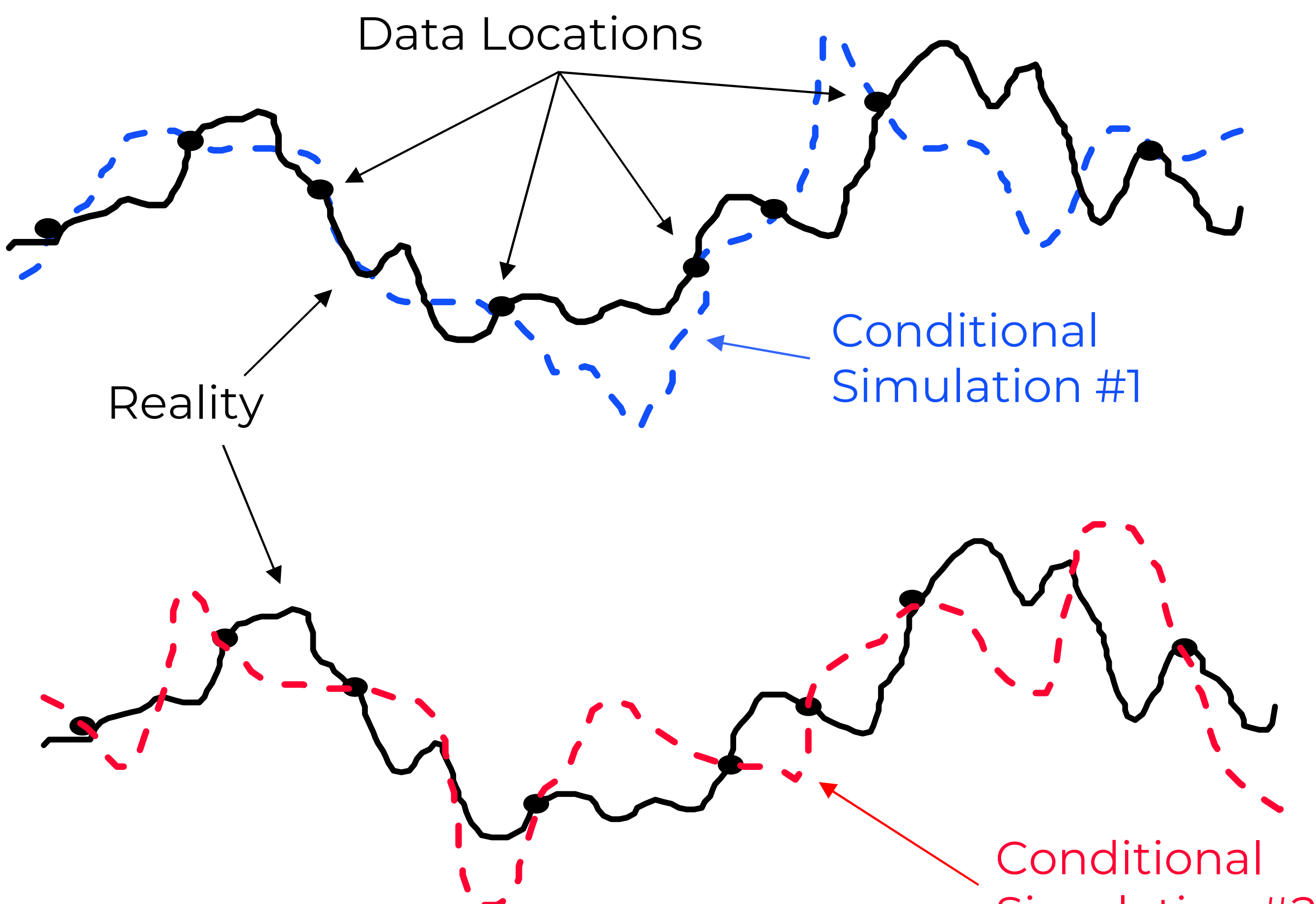
Recall estimation: assign the most accurate value at each location



- Estimation is jointly wrong
  - The spatial continuity is too high
  - The variance is too low
  - We need a method that corrects for the smoothing simulation

MOTIVATION FOR SIMULATION

Simulation sacrifices local accuracy for global accuracy and provides an uncertainty model with multiple realizations



MOTIVATION FOR SIMULATION

Simulation generates models that mimic the target phenomenon

What does 'simulated dill pickle' potato chips taste like?

What does a simulated reservoir model look like?

What does a simulated reservoir flow like?

ESTIMATION VS. SIMULATION

Simulation generates models that mimic the target phenomenon

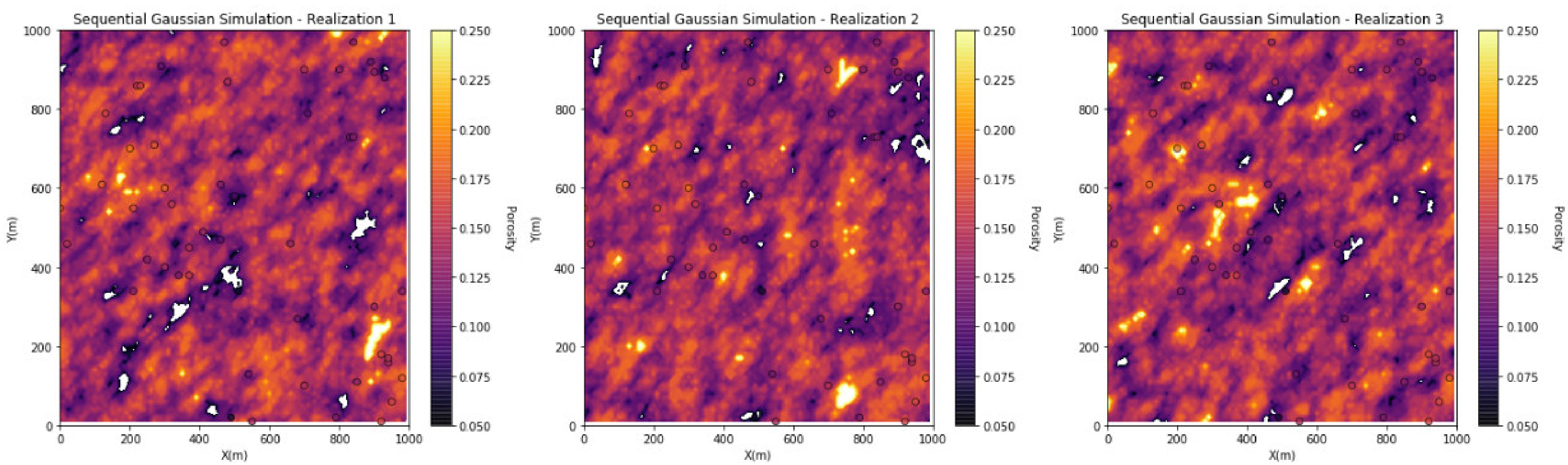
- Estimation:
  - honors local data (with discontinuity)
  - locally accurate
  - smooth appropriate for visualizing trends
  - inappropriate for flow simulation
  - no assessment of global uncertainty

- Simulation:
  - honors local data
  - reproduces histogram
  - honors spatial variability → appropriate for flow simulation
  - alternative realizations possible → change random number seed
  - assessment of global uncertainty is possible

SIMULATION EXAMPLE

Multiple Realizations

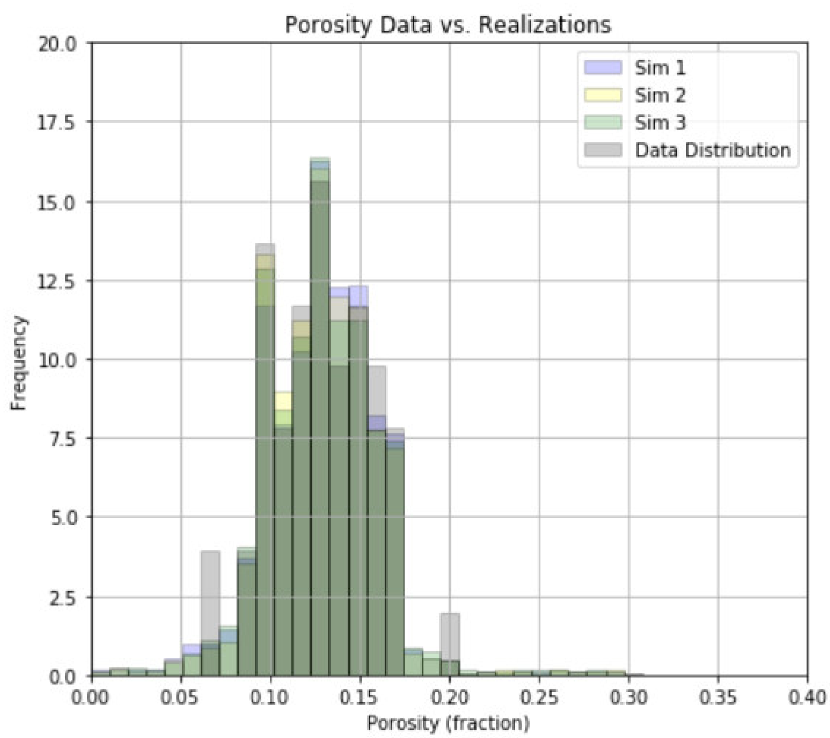
- Sequential Gaussian Simulation (Univariate) – 3 Realizations
  - Observed the behavior at and away from data over the three realizations
  - Observe the spatial continuity

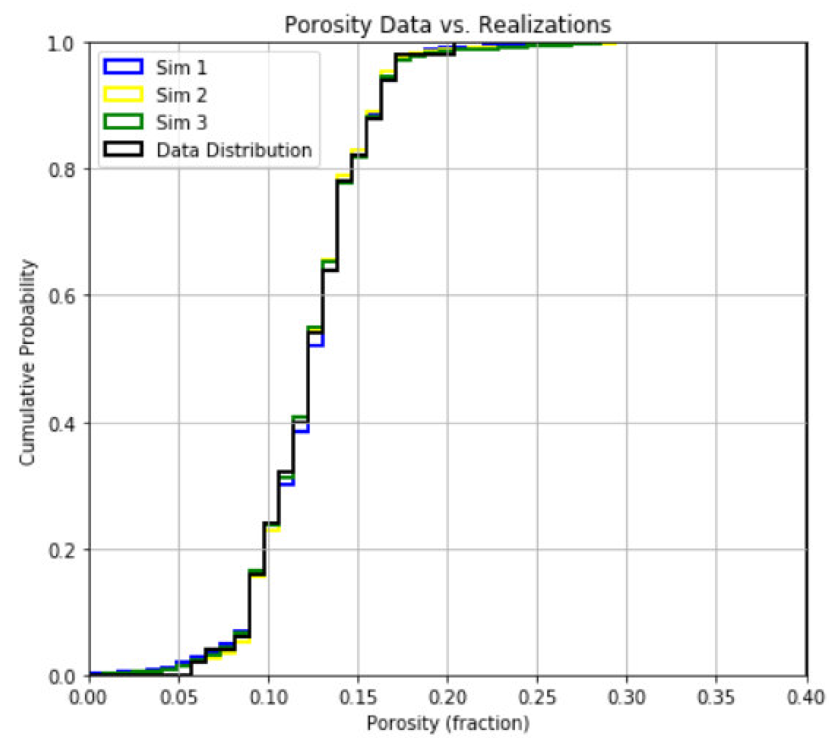


SIMULATION EXAMPLE

Multiple Realizations

- Sequential Gaussian Simulation (Univariate)
  - check the realization distributions

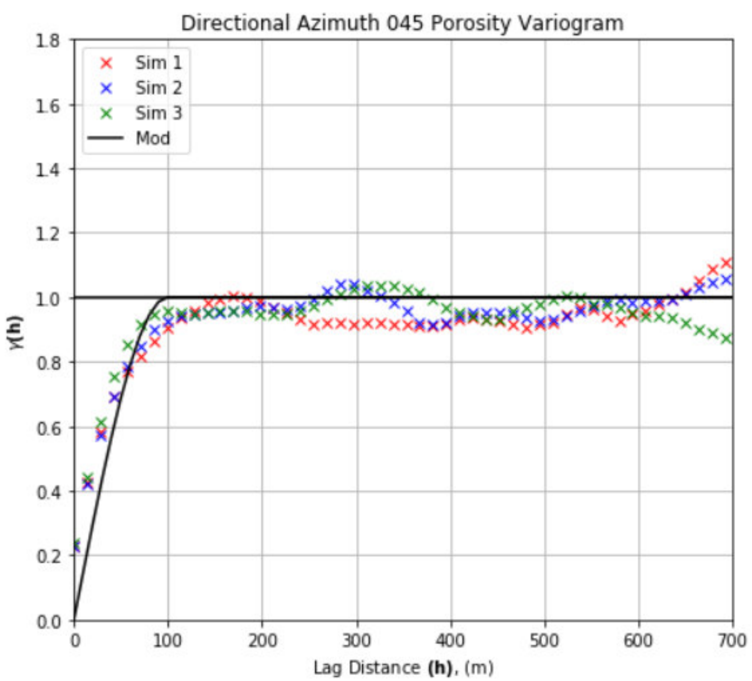


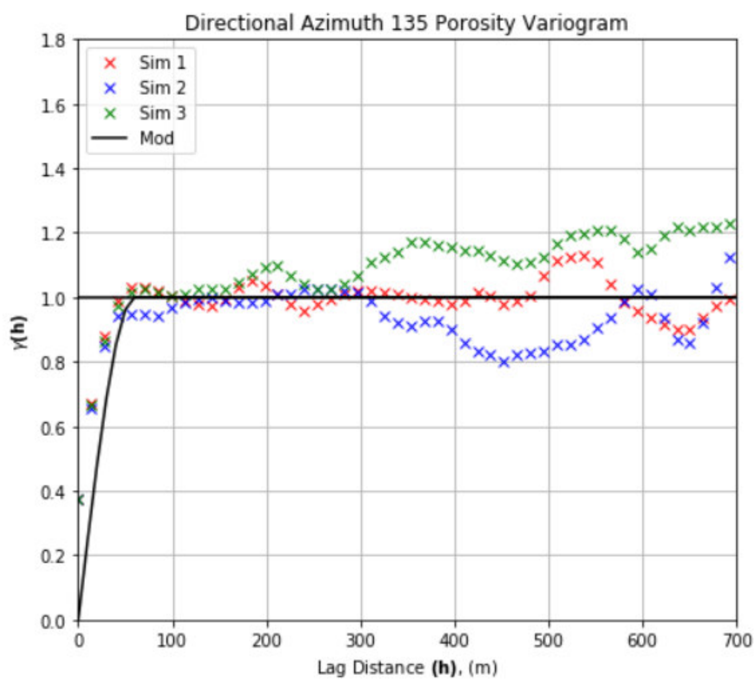


SIMULATION EXAMPLE

Multiple Realizations

- Sequential Gaussian Simulation (Univariate)
  - check the spatial continuity, variograms





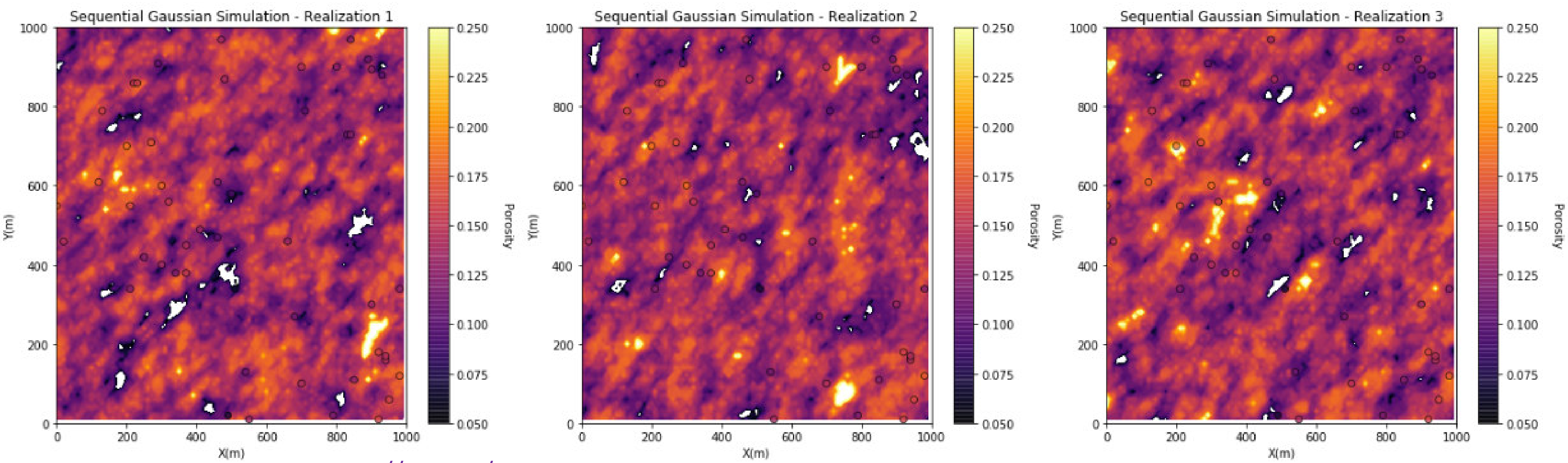
SIMULATION HANDS-ON

Multiple Realizations

- Sequential Gaussian Simulation (Univariate)

Things to try:

1. Adjust the variogram parameters and rerun.
2. Azimuth, anisotropy ratio and type of variogram model structure
3. Switch between simple kriging and ordinary kriging.



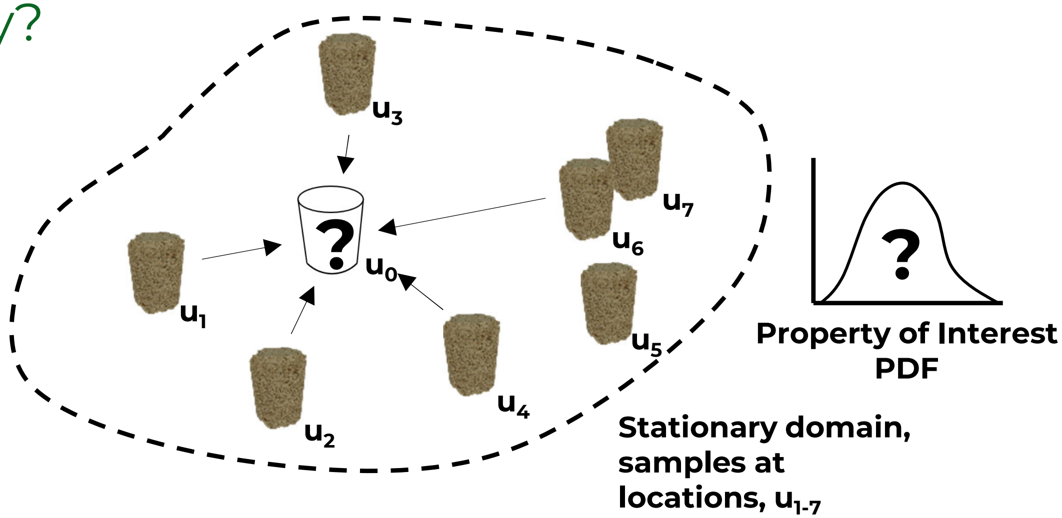
The file is at: https://git.io/fjTpu

The file is GeostatsPy_univariate_simulation.ipynb

SOURCES OF UNCERTAINTY

UNCERTAINTY

What is uncertainty?



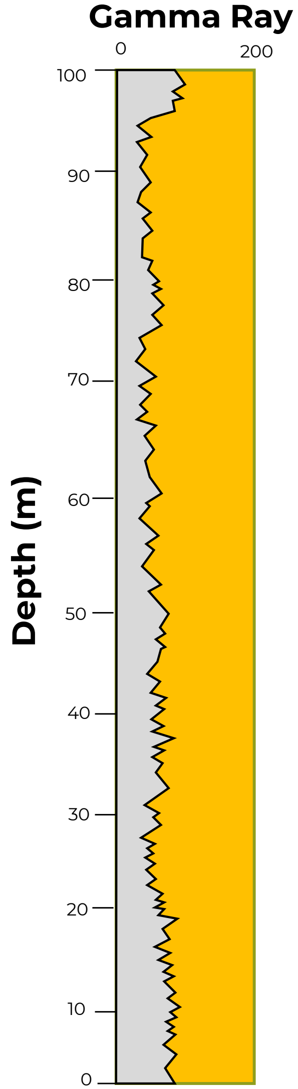
- Uncertainty is not an intrinsic property of the subsurface.
  - At every location ($\mathbf{u}_\alpha$) within the volume of interest the true properties could be measured if we had access (facies, porosity etc.).
  - **Uncertainty is a function of our ignorance**, our inability to observed and measure the subsurface with the coverage and scale required to support our scientific questions and decision making.

sparsity of sample data + heterogeneity = uncertainty

  - If the subsurface was homogeneous, with a few measurements, uncertainty would be reduced and estimates resolved to a sufficient degree of exactitude.

TYPES OF UNCERTAINTY

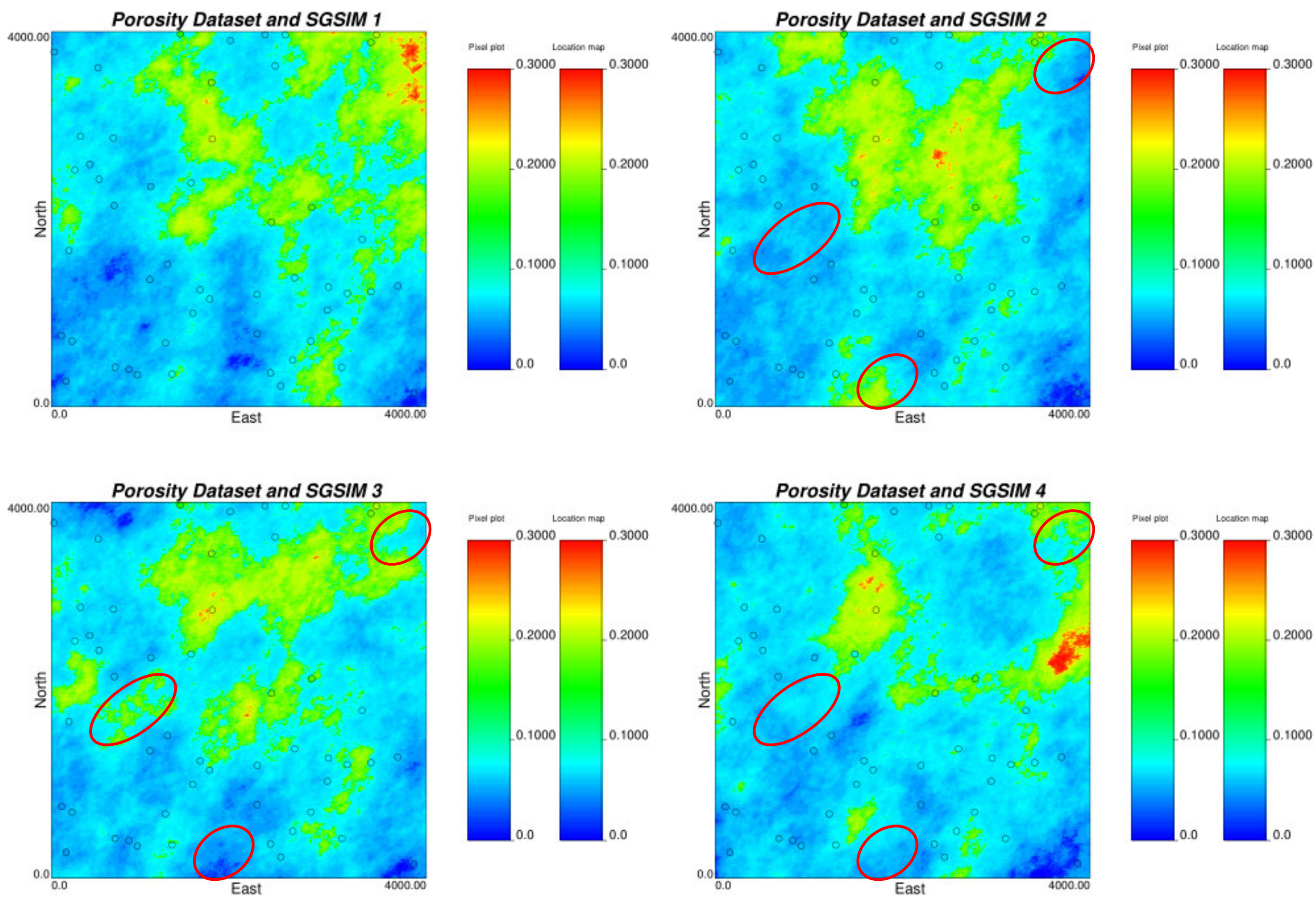
- Measurement / Interpretation Error
  - Formation evaluation – tool tolerance, calibration error, approximations / assumptions
  - Interpreter experience and prior model / assumptions
  - How to integrate it?
    - Indicator method code as soft inputs
    - Multiple data realizations in design of experiments



Example well log.

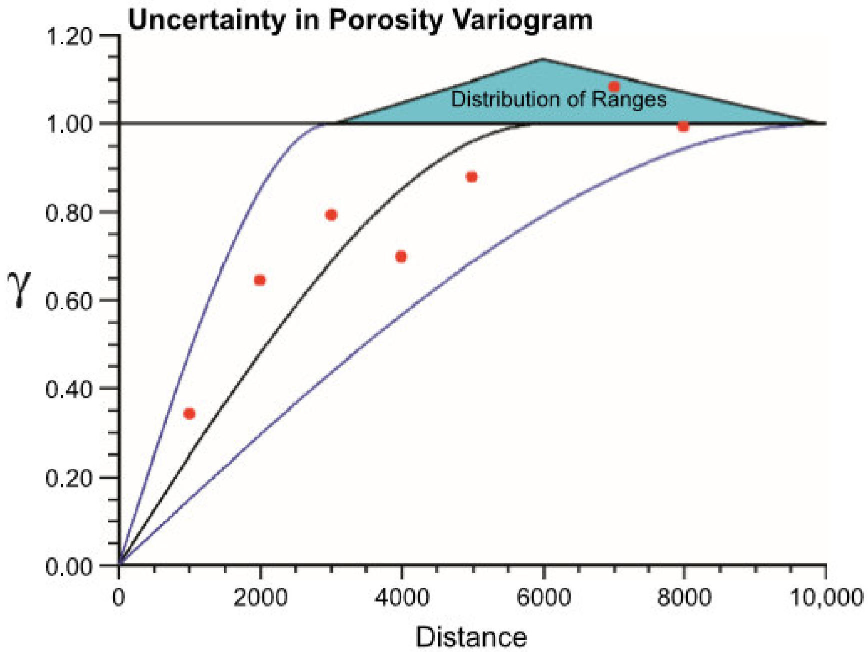
TYPES OF UNCERTAINTY

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

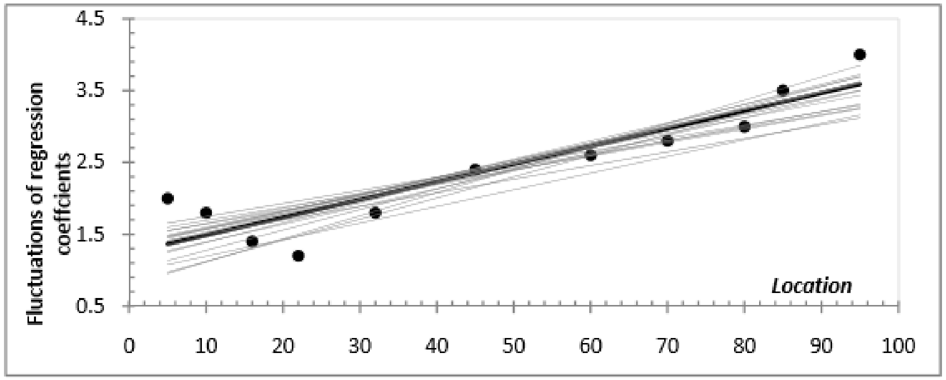


TYPES OF UNCERTAINTY

- Parameter Uncertainty
  - Uncertainty in the input statistics to constrain the model area
  - E.g. global reference porosity distribution for simulation
  - Formulate distribution scenarios and bootstrap realizations



Distribution of Variogram Ranges (Pyrcz et al., 2006)



Trend Uncertainty (Villalba and Deutsch., 2010)

REPRESENTING UNCERTAINTY

HOW DO WE REPRESENT UNCERTAINTY?

- **Using Multiple Models**: We represent uncertainty with multiple models.

- **Scenarios**: when the input decisions and parameters are changed

  Captures interpretation and data uncertainty.

- **Realizations**: when the input decisions and parameters are held constant and only the random number seed is changed

  Captures spatial uncertainty.

- **Working With Multiple Models**: It is generally not appropriate to analyze a single or few scenarios and realizations.

- **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.).

HOW DO WE REPRESENT UNCERTAINTY?

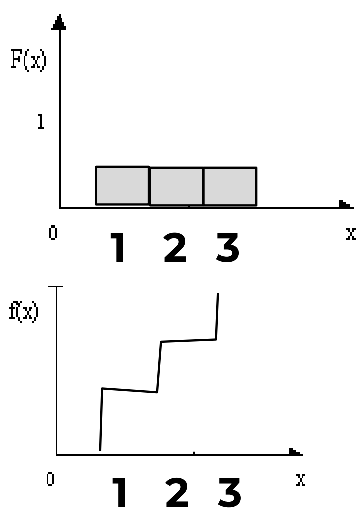
- We have a PDF / CDF of a measure over a volume: therefore our measures of spread / dispersion are our measures of uncertainty.

- Variance: Var$(Z) = \int_{-\infty}^{\infty}(z-m)^2 f(z)dz$ **(Expected squared difference from mean)**

- Dispersion Variance: $D^2(v,V) = \bar{\gamma}_{V,V} - \bar{\gamma}_{v,v}$ **(Generalized variance accounting scale and heterogeneity)**

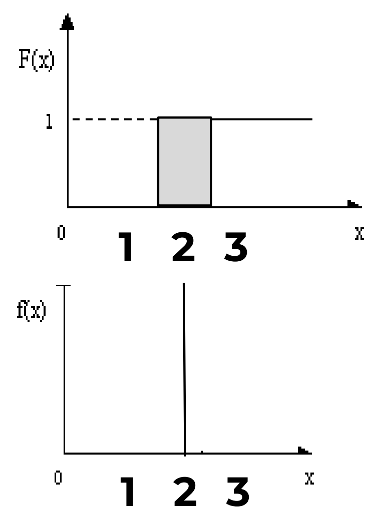
- Entropy: $H(Z) = -\sum_p^n P(Z_i) \cdot ln\, P(Z_i)$ **(Measure of uncertainty for categorical variables)**

Discrete distribution



maximum entropy

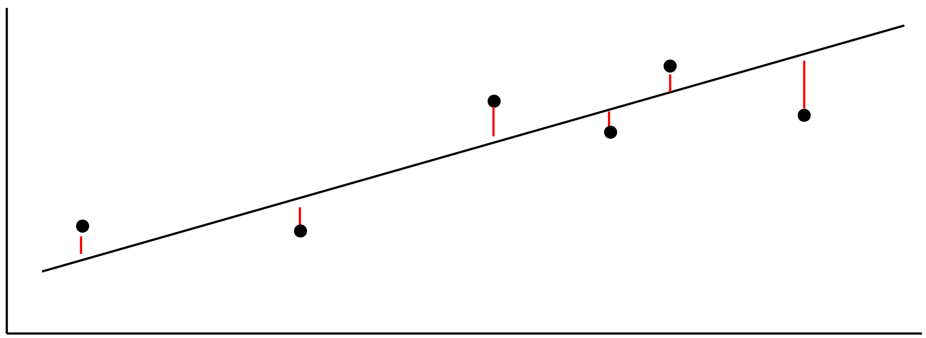
Discrete distribution



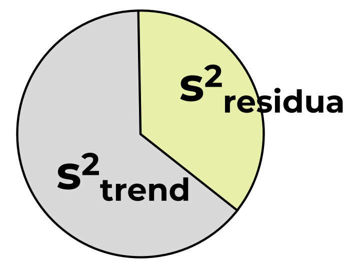
minimum entropy

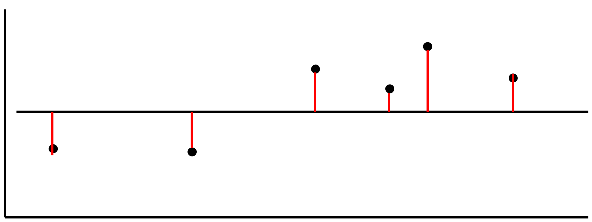
HOW DO WE REPRESENT UNCERTAINTY?

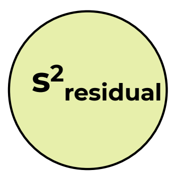
- Variance is partitioned between trend, deterministic, known and residual, stochastic, unknown.
- Variance Components: $\sigma^2 = \sigma_t^2 + \cancelto{0}{\sigma_r^2} + 2\,C_{t,r}(0)$
- Total variance = Deterministic / Known Variance + Stochastic / Unknown Variance

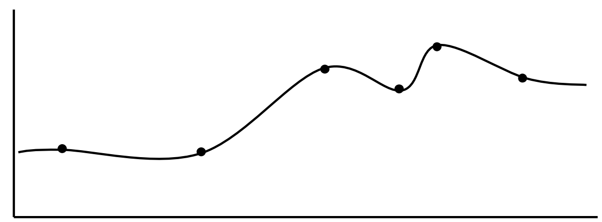


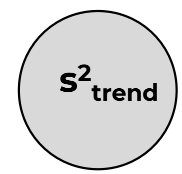
Variance Partitions











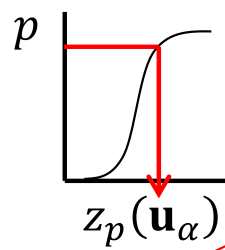
SUMMARIZING OVER REALIZATIONS

SUMMARIZING UNCERTAINTY OVER MULTIPLE REALIZATIONS

- Multiple Realizations
  - Visualizing / Communicating Uncertainty
  - We need practical workflows to summarize over multiple realizations
  - Local uncertainty maps provide measures of local uncertainty suitable to support decision making (more on this later)

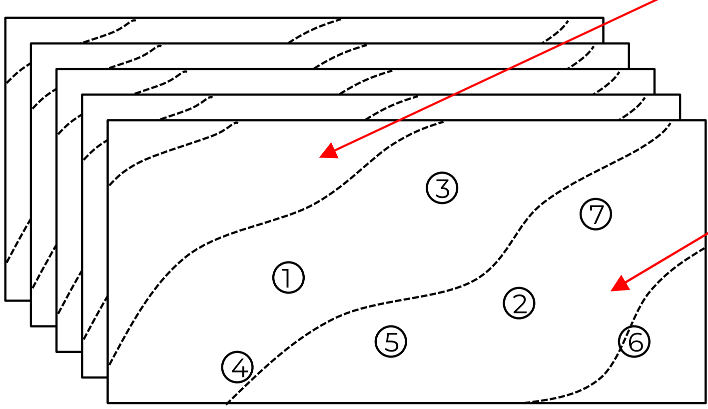
SUMMARIZING UNCERTAINTY OVER MULTIPLE REALIZATIONS

- Method:
  - We need to work with multiple models, how do we summarize?
  - Scan over all the realizations and scenarios
  - Calculate the local distributions of uncertainty at each location
  - Calculate statistical summary over each location and place in a map / model



What is a specific percentile outcome at this location?

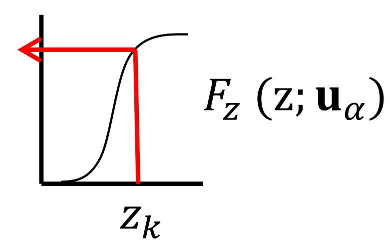
$$z_p(\mathbf{u}_\alpha) = F_Z^{-1}(p; \mathbf{u}_\alpha)$$



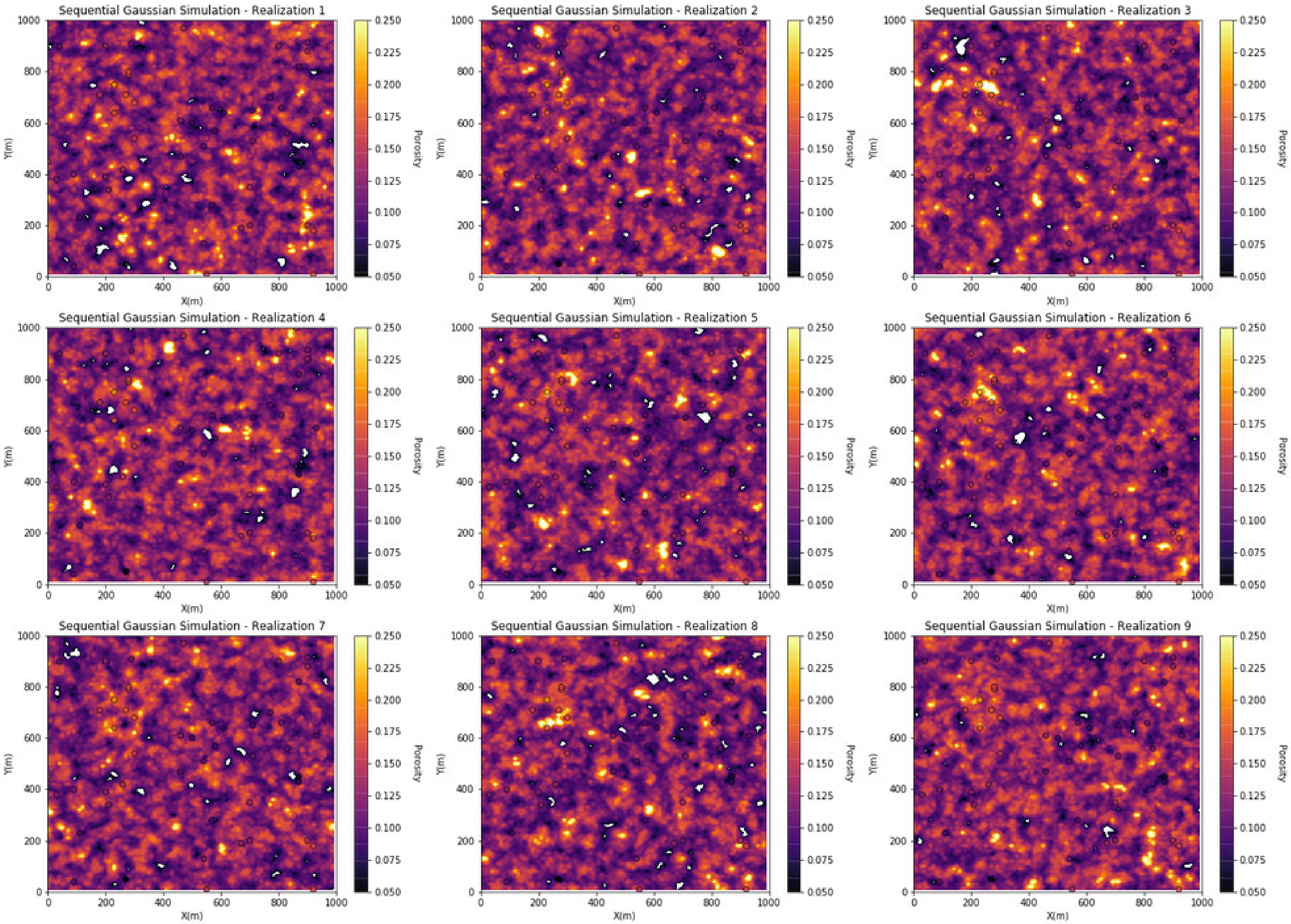
What is the local probability of exceeding a threshold at this location?

$$P(\mathbf{z}(\mathbf{u}_\alpha) > z_k) = 1 - F_z\,(z_k; \mathbf{u}_\alpha)$$

$$P(\mathbf{z}(\mathbf{u}_\alpha) < z_k)$$

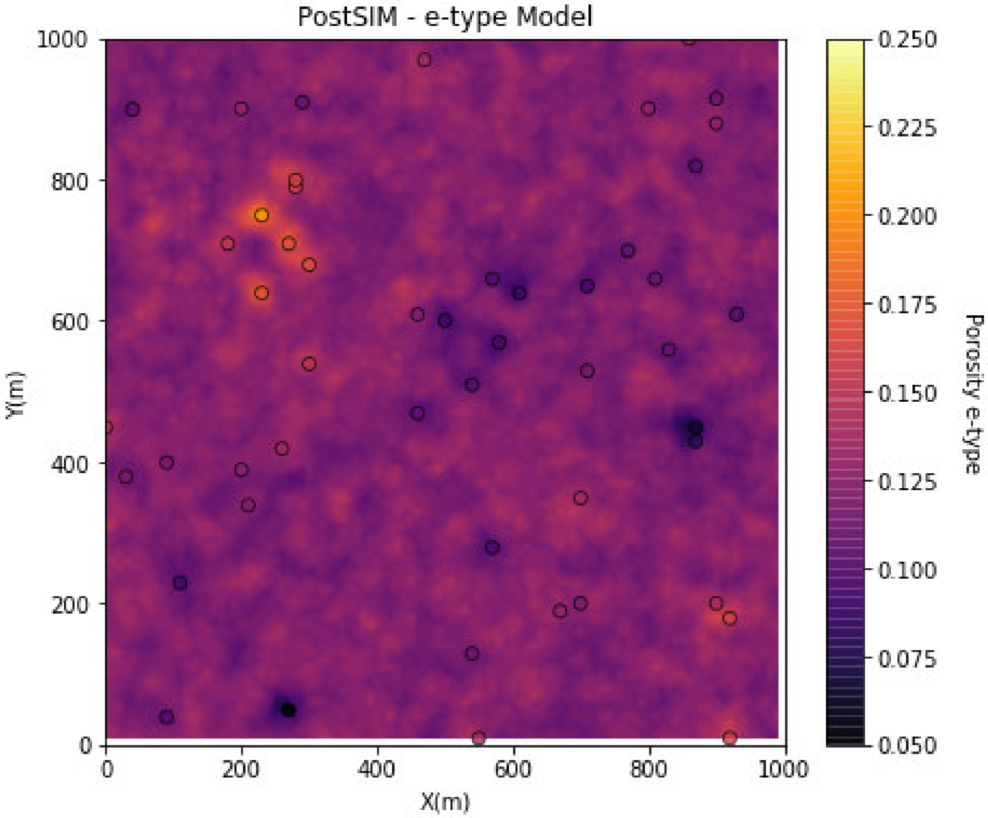


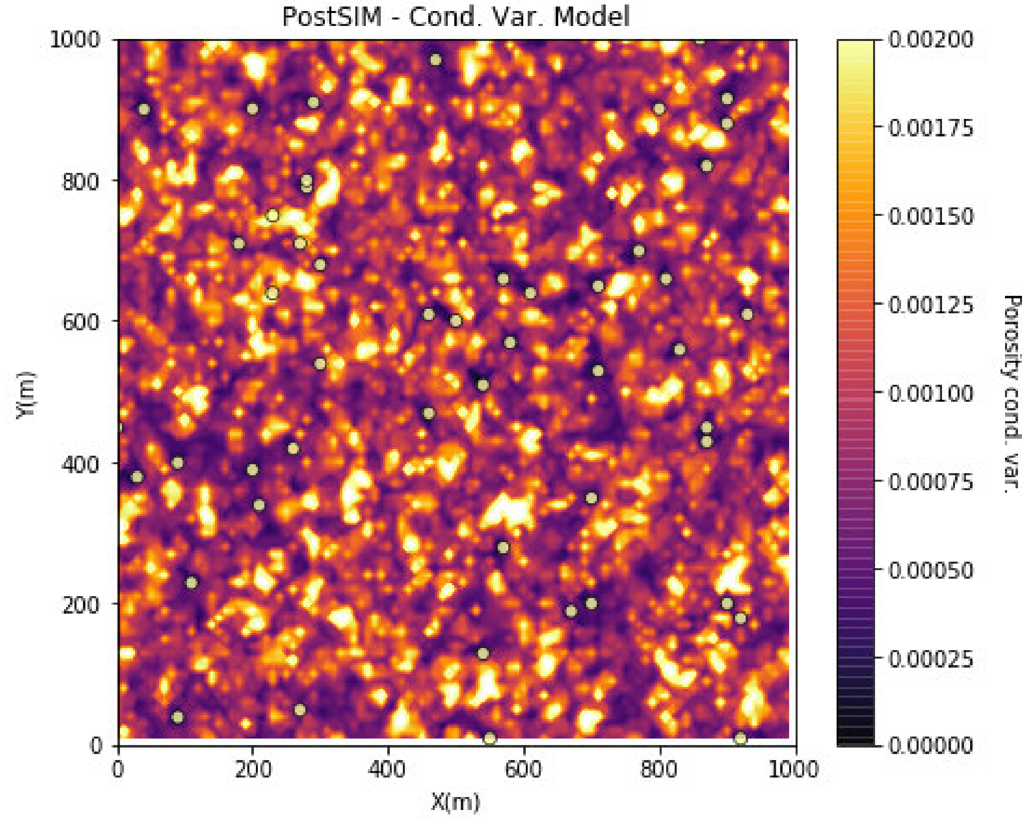
LOCAL UNCERTAINTY EXAMPLE



9 Realizations of porosity

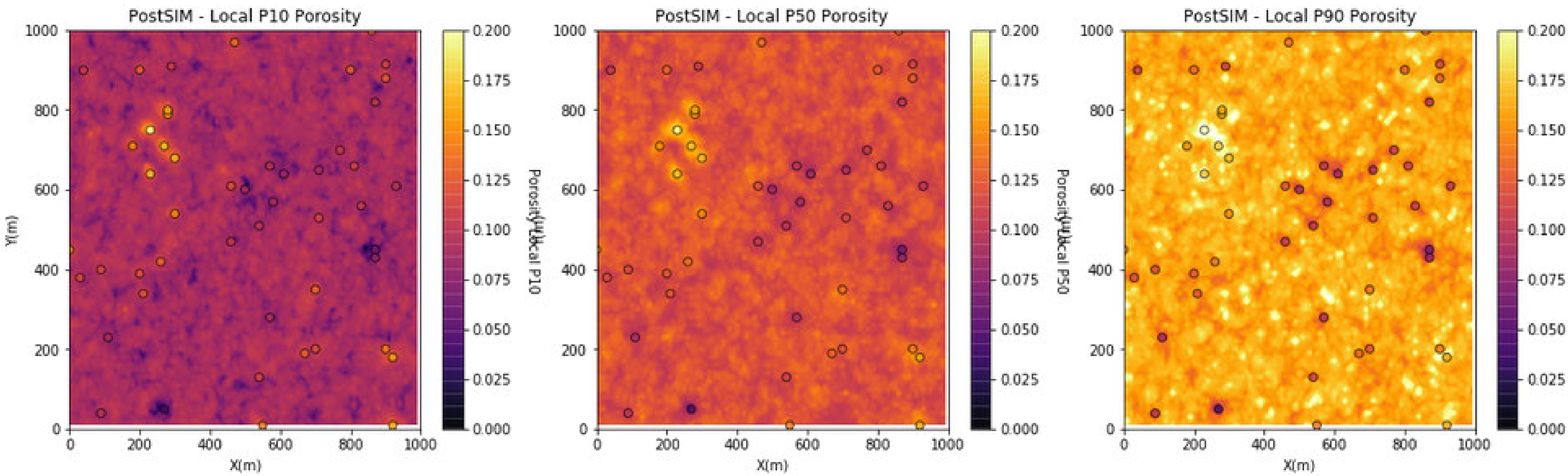
LOCAL UNCERTAINTY EXAMPLE





- We will start with the e-type and the conditional variance.
  - e-type is the local expectation, the average of the L realizations at location $\mathbf{u}_\alpha$ as we assume all realizations are equally likely
  - conditional variance is the local variance

LOCAL UNCERTAINTY EXAMPLE



- Local percentile maps are the maps with the local percentile values sampled from the local realizations
- We can interpret these as follows, at a location if we have a local P10 of 14% porosity, then we have a 90% probability of an even higher porosity, the porosity at that location is surely high
- Local percentiles are very convenient to understand local uncertainty. We must NOT confuse them with a percentile model (the model that is globally ranked as a specific percentile outcome)

LOCAL UNCERTAINTY HANDS-ON

- Realization Post-processing

- Things to try:

1. Observed the results at and away from the data
2. Adjust the percentiles
3. Adjust the thresholds

The file is:

Daytum_Simulation_Postsim.ipynb

Daytum +2 Course: Data Analytics, Geostatistics and Machine Learning Deep Dive

Simulation Post-processing Demonstration and Exercise

Goal

Calculate local uncertainty over an ensemble of spatial models.

Description

Here's a simple, documented workflow, demonstration of local uncertainty calculation for subsurface modeling workflows. This should help you get started with building subsurface models that integrate spatial continuity.

Here's a simple workflow for realization post-processing. Some examples of applications include:

- Quantifying local uncertainty away from wells
- Accessing the probability / risk of specific local outcomes

First let's explain the concept of realization post-processing.

Realization Post-processing

Post-processing refers to operations to provide summaries over multiple realizations.

Here will will focus on a few local statistical summaries. These methods calculate the local cumulative distribution function at each location in the model based on pooling the local realizations.

$$F_x(\mathbf{u}_\alpha)$$

The following are local summaries demonstrated in this workflow:

- **e-type** is the local expectation (since equal weighted the same as the average)

UNCERTAINTY SUMMARY

UNCERTAINTY COMMENTS

- **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. Overfit 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.

- **Be an uncertainty detective! Discover and evaluate various sources**

UNCERTAINTY COMMENTS

- **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.

- **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.

- **Ignoring Uncertainty is Assuming Certainty**

UNCERTAINTY NEW TOOLS

Topic	Application to Subsurface Modeling
Simulation	Correct the smoothing of kriging and represent uncertainty through multiple realizations. *Use simulation to capture data measurement, interpretation and spatial uncertainty.*
Sources of Uncertainty	Seek out and integrate all significant sources of uncertainty. *Uncertainty in the data measures based on data realizations combined with spatial uncertainty with multiple spatial realizations.*
Scenarios	Capture uncertainty in model parameters and decisions through multiple scenarios. *Construction of 3 set of realizations for low, mid and high case spatial continuity models.*
Calculating Uncertainty	Use bootstrap, model resampling, expert assessments, Bayesian updating etc. *Use a combination of expert judgement, analogs and robust statistical methods.*

Uncertainty

Lecture outline . . .

- Stochastic Simulation
- Sources of Uncertainty
- Representing Uncertainty
- Summarizing Over Realizations
- Uncertainty Summary