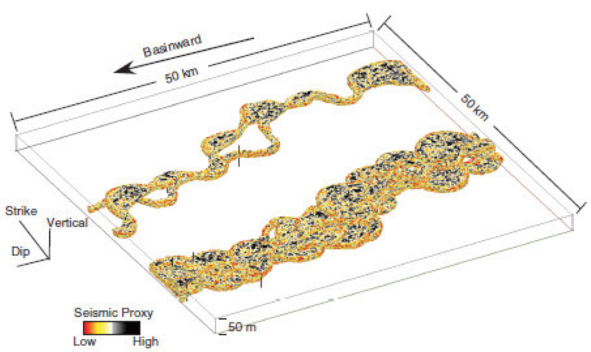
Co-simulation

Lecture outline . . .

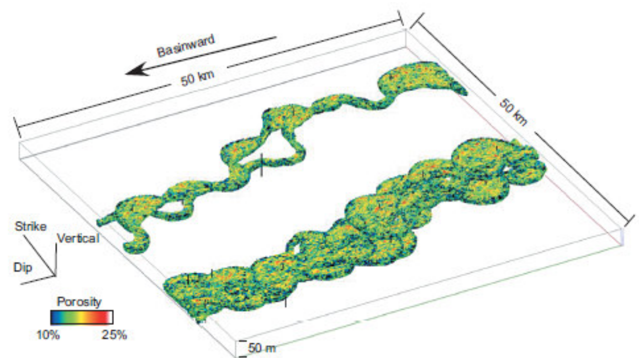
- Co-simulation
- Full Cokriging
- Collocated Cokriging
- P-field Simulation

MOTIVATION

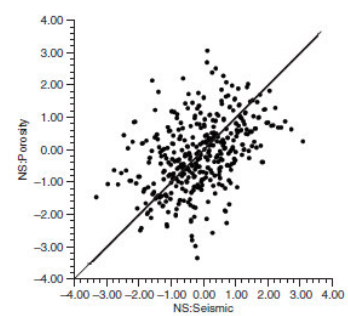
- We need subsurface modeling methods that:
  - Account for the relationships between multiple correlated variables, e.g., seismic acoustic impedance and porosity.



Secondary Realization



Primary Realization



Bivariate Realationship

CO-SIMULATION

COMMON MODELING WORKFLOW

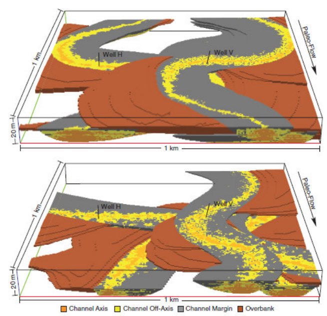
Facies Categorical, Porosity Continuous then Permeability Co-simulation – here's Some Context!

Categorical Simulation

Models Facies PDF, Trends and Spatial Continuity

→

Multiple Facies Realizations



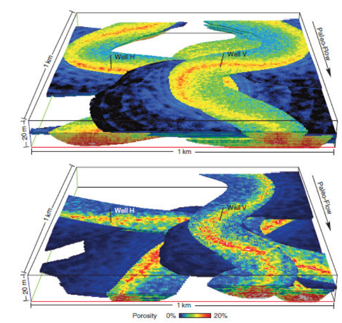
→

Continuous Simulation

Models Porosity PDF, Trends and Spatial Continuity within each facies

→

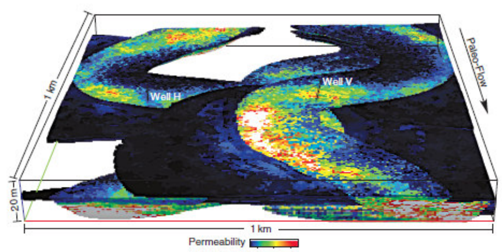
Multiple Porosity By-facies Realizations



Continuous Cosimulation

Models Permeability PDF, Trends and Spatial Continuity and Porosity Relationship within each facies

Multiple Permeability By-facies Realizations Correlated with Porosity



Note: only had 1 realization on hand (should be two in figure).

The common geostatistical reservoir modeling workflow, co-simulation step in red box.

COMMON MODELING WORKFLOW

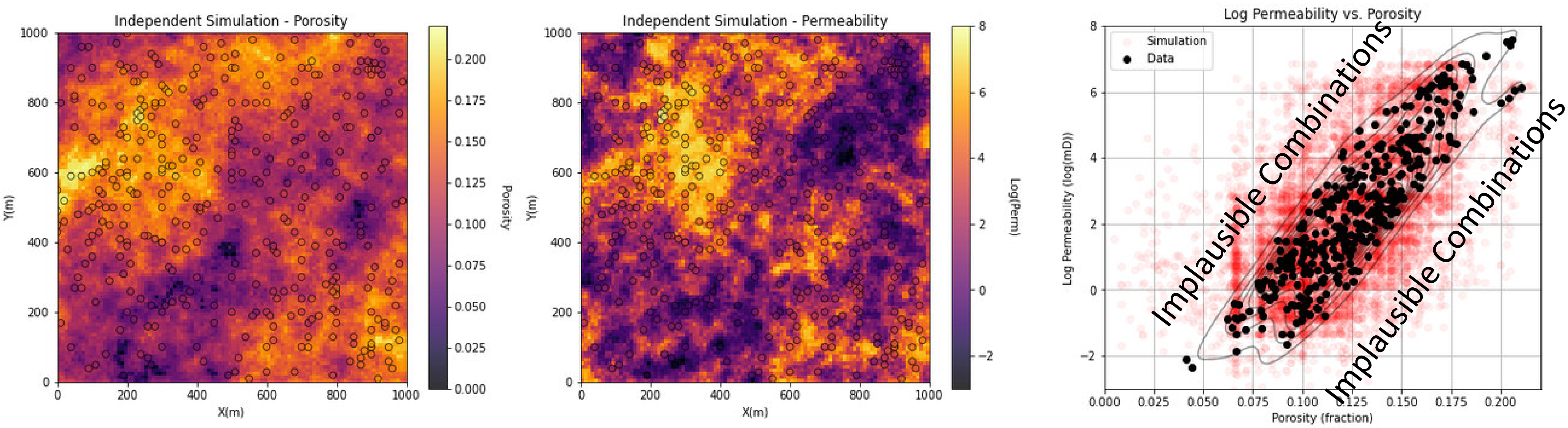
- We typically need to build reservoir models of more than one property of interest.

- For example, a typical workflow is:
  - Build facies models realizations
  - Within the facies realizations simulate the porosity
  - Within the facies realizations co-simulate the permeability (primary variable) correlated with the previously simulated porosity realization (secondary variable)

- We will only cover the commonly applied co-simulations methods
  - collocated cokriging and cloud transform
  - we limit ourselves to simulating one property correlated to one other (bivariate)

CO-SIMULATION

- Co-simulation is not perfect, but the alternative is to simulate each property independently
  - no correlation or some minor degree correlation due to data conditioning (correlation at the data locations gets propagated away from the data in each feature).

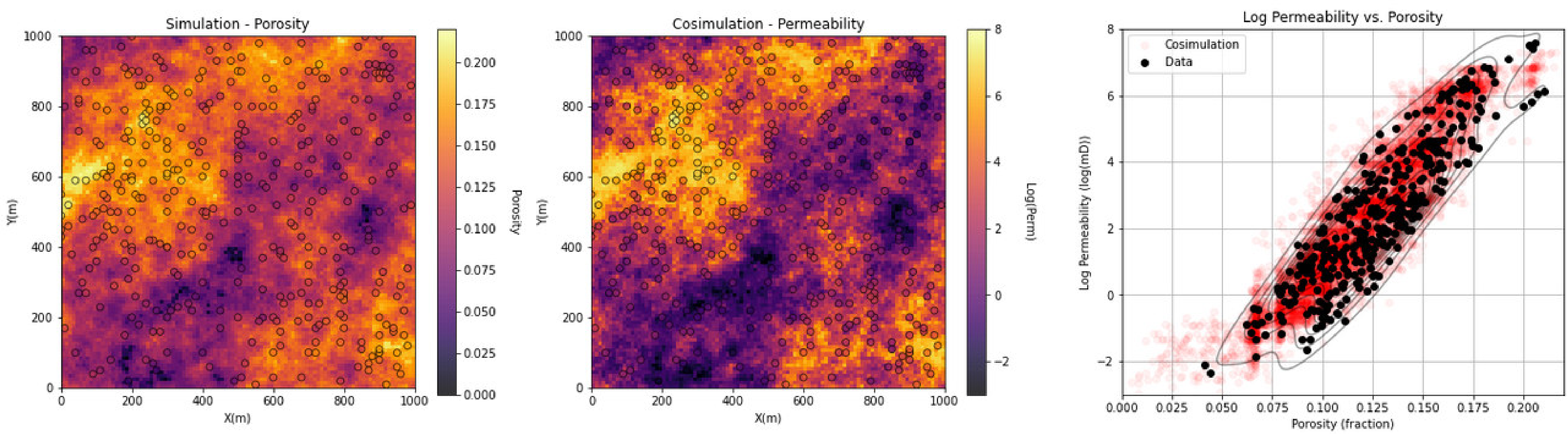


Independent simulation of porosity (left) and permeability (center) and cross plot data and simulated values (right), file is GeostatsPy_cosimulation.ipynb.

- This results in implausible combinations away from data and too high uncertainty, because we ignore information between the features.

CO-SIMULATION

- Now we repeat the previous example with co-simulation.
  - an example with good reproduction of bivariate features, correlation based on collocated cokriging

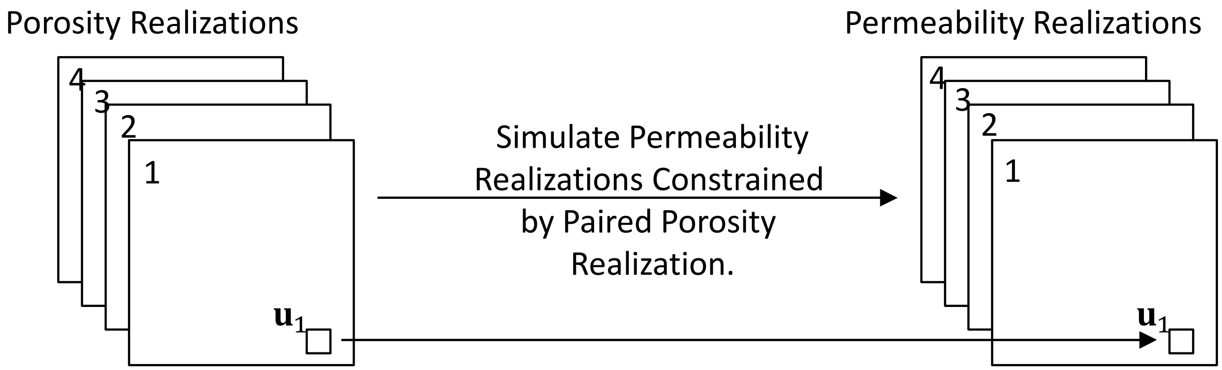


Co-simulation of permeability (center) given porosity (left) and cross plot data and simulated values (right), file is GeostatsPy_cosimulation.ipynb.

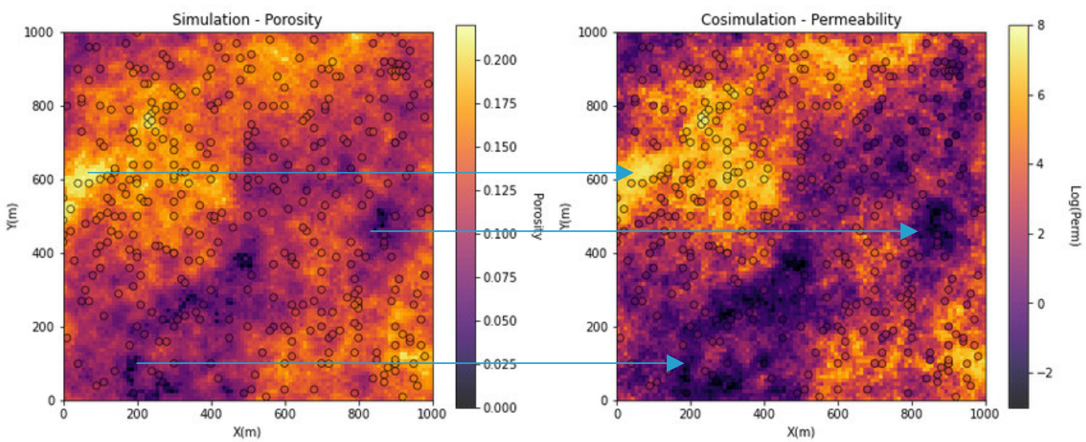
- Capture the general correlation structure with some tail extrapolation.

CO-SIMULATION

- Co-simulation is a simulation method that imposes correlation with a previously simulated property.
  - The realizations are paired.



- Honor the bivariate relationship between pair values at the same locations, $\mathbf{u}_\alpha$



Co-simulation of permeability (center) given porosity (left) and cross plot data.

CO-SIMULATION

- Each co-simulation method will have a 'conditioning' priority:
  - **Collocated Cokriging** prioritizes the histogram and variogram and may honor the correlation coefficient between the two variables
  - **Cloud transform** will honor the specific form of the bivariate relationship (cloud) between the two features but may not honor the histogram nor the variogram.
- These methods start with a completed realization of the secondary feature
  - e.g., porosity secondary feature, if we are co-simulated permeability primary feature constrained by porosity
- Multiple information sources may be contradictory, in this cases lower priority information is preferentially sacrificed

CO-SIMULATION

- Commonly used term in geostatistics for constraining a geostatistical model to honor / reproduce data, statistics and interpretations (trends, stationary domains) derived from data and / or expert knowledge.

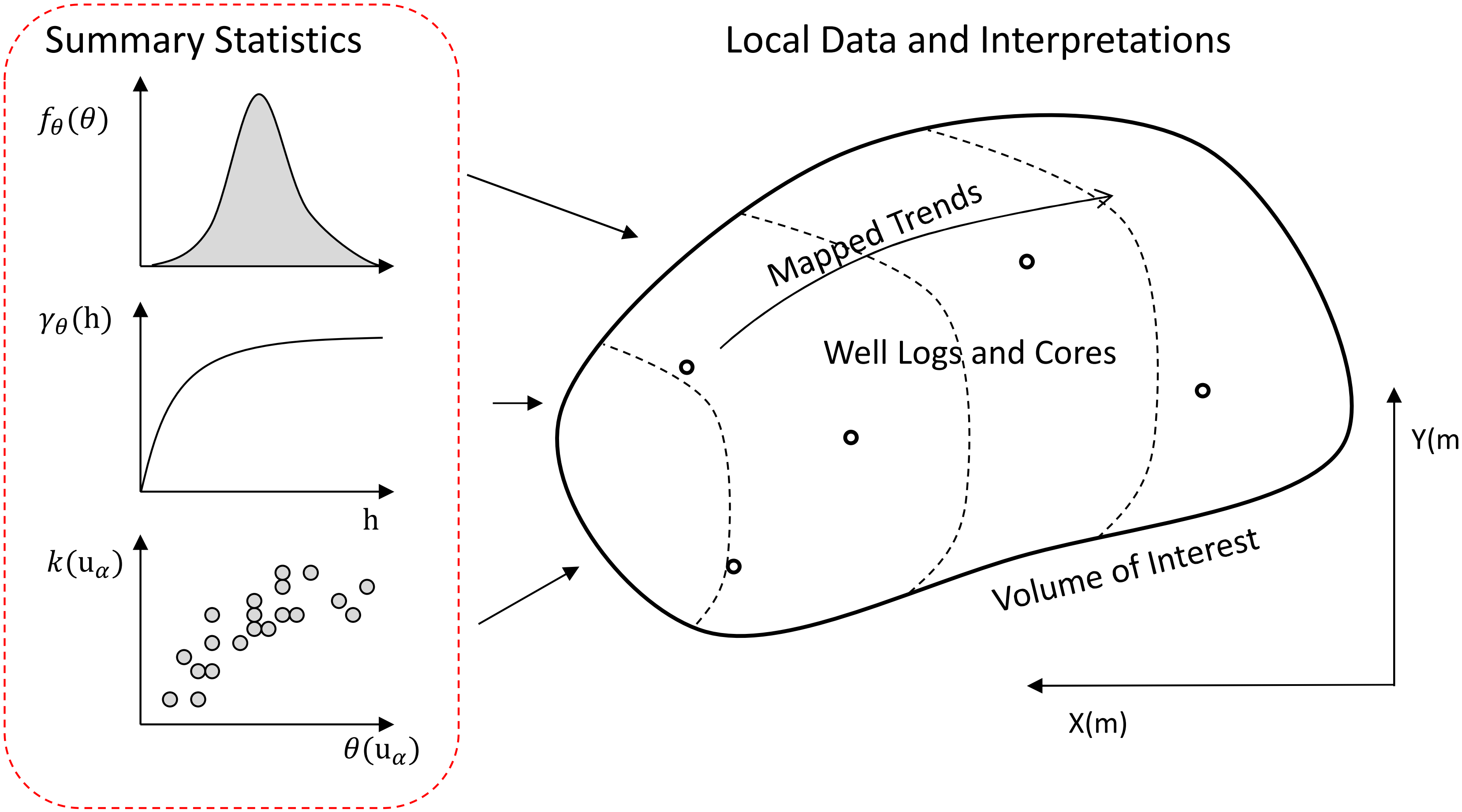
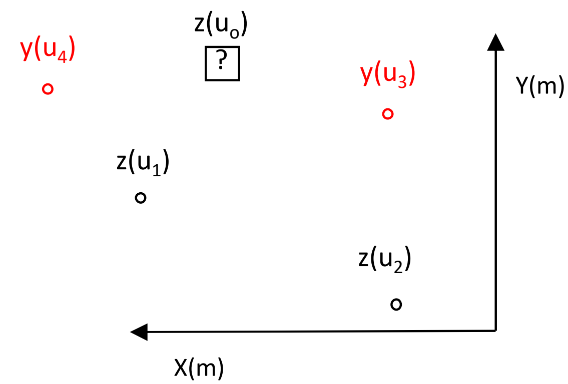


Illustration of various types of constraints to condition a geostatistical model.

FULL COKRIGING

FULL COKRIGING

- We can extend the simple kriging system to integrate other features.
- For example, see this data (right) and the cokriging system (below):
  - primary feature at $\mathbf{u}_1$ and $\mathbf{u}_2$
  - secondary feature at $\mathbf{u}_3$ and $\mathbf{u}_4$
  - to estimate primary feature at $\mathbf{u}_0$



Data (8) and estimate (?) locations.

Redundancy | Weights | Closeness

Direct: Primary, Primary | Cross: Primary, Secondary | Primary | Direct: Primary

$$
\begin{bmatrix}
C_z(\mathbf{u}_1,\mathbf{u}_1) & C_z(\mathbf{u}_1,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_3) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_4) \\
C_z(\mathbf{u}_2,\mathbf{u}_1) & C_z(\mathbf{u}_2,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_2,\mathbf{u}_3) & C_{z,y}(\mathbf{u}_2,\mathbf{u}_4) \\
C_{z,y}(\mathbf{u}_3,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_3,\mathbf{u}_2) & C_{y,y}(\mathbf{u}_3,\mathbf{u}_3) & C_y(\mathbf{u}_3,\mathbf{u}_4) \\
C_{z,y}(\mathbf{u}_4,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_4,\mathbf{u}_2) & C_{y,y}(\mathbf{u}_4,\mathbf{u}_3) & C_y(\mathbf{u}_4,\mathbf{u}_4)
\end{bmatrix}
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix}
=
\begin{bmatrix}
C_z(\mathbf{u}_o,\mathbf{u}_1) \\
C_z(\mathbf{u}_o,\mathbf{u}_2) \\
C_{z,y}(\mathbf{u}_o,\mathbf{u}_3) \\
C_{z,y}(\mathbf{u}_o,\mathbf{u}_4)
\end{bmatrix}
$$

Cross: Secondary, Primary | Direct: Secondary, Secondary | Secondary | Direct: Secondary

Full cokriging system to 2 primary data and 2 secondary data.

FULL COKRIGING

- **Cross Variogram**, measure of how two variables differ together over distance.

$$\gamma_{z,y}(\mathbf{h}) = \frac{1}{2} E\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})][Y(\mathbf{u}) - Y(\mathbf{u}+\mathbf{h})]\}$$

- **Cross Covariance**, measure of how two variables vary together over distance.

$$C_{z,y}(\mathbf{h}) = E\big\{[Z(\mathbf{u}) - m_z]\big[Y(\mathbf{u}+\mathbf{h}) - m_y\big]\big\}$$

$$= E\{[Z(\mathbf{u})][Y(\mathbf{u}+\mathbf{h})]\} - m_z \cdot m_y \quad \forall\, \mathbf{u}$$

above can be shown with expectation.

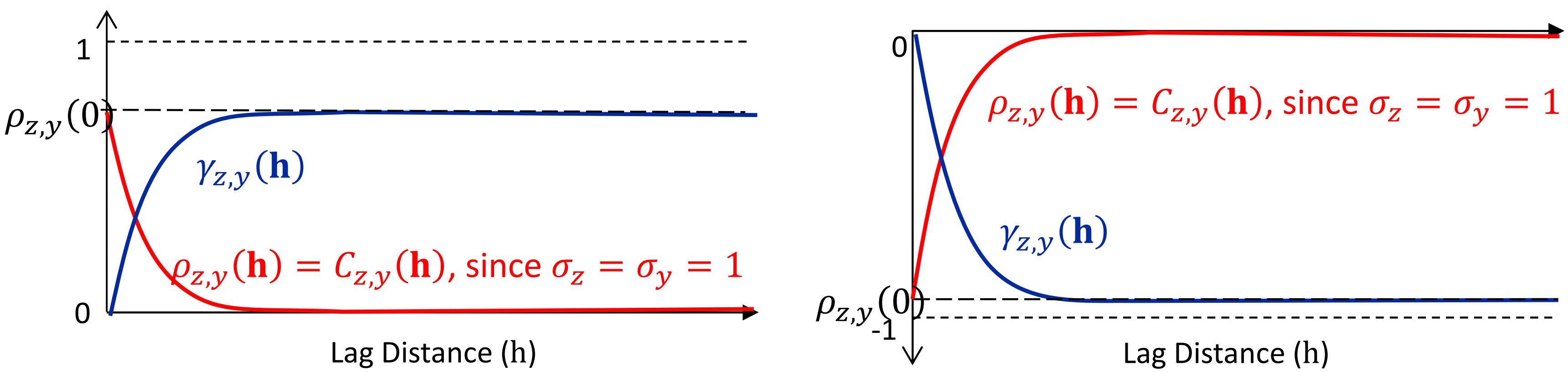
- **Cross Correlogram** is the standardized cross covariance. Correlation coefficient vs. lag distance!

$$\rho_{z,y}(\mathbf{h}) = \frac{C_{z,y}(\mathbf{h})}{\sigma_z \cdot \sigma_y} \qquad \rho_{z,y}(\mathbf{h}) = C_{z,y}(\mathbf{h}), \text{ if } \sigma_z = \sigma_y = 1.0$$

FULL COKRIGING

- Cross variogram starts at 0.0, $\gamma_{z,y}(\mathbf{0}) = \mathbf{0}$, and then at the range reaches the correlation coefficient, $\gamma_{z,y}(\mathbf{h}) \rightarrow \rho_{z,y}(0)$, as $\mathbf{h}$→ range.
  - if the correlation coefficient is less than zero, $\rho\_(z,y)$ (0)<0, then the cross variogram is negative! and sill is the correlation coefficient

- Cross correlogram (equal to cross covariance if $\sigma_z = \sigma_y = 1$), starts at the correlation coefficient, $\rho_{z,y}(\mathbf{0}) = \rho_{z,y}$, and then at the range reaches the 0, $\rho_{z,y}(\mathbf{h}) \rightarrow \mathbf{0}$, as $\mathbf{h}$→ range.
  - if the correlation coefficient is less than zero, $\rho_{z,y}(0) < 0$, then the cross correlogram is negative!



Cross variogram and cross correlograms for features Y and Z with positive correlation (left) and negative correlation (right).

COLLOCATED COKRIGING

COLLOCATED COKRIGING

- Collocated Cokriging makes two simplifications of full cokriging:
  - Only one (the collocated) secondary variable is considered
  - Cross covariance $C_{z,y}(\mathbf{h})$ is assumed to be a linear scaling of $C_z(\mathbf{h})$

- The collocated secondary value is surely the most important and likely screens the influence of multiple secondary data
  - Consider the implications for the cokriging system with only the collocated secondary data value included (below).

$$\begin{bmatrix} C_z(\mathbf{u}_1,\mathbf{u}_1) & C_z(\mathbf{u}_1,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_3) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_4) \\ C_z(\mathbf{u}_2,\mathbf{u}_1) & C_z(\mathbf{u}_2,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_2,\mathbf{u}_3) & C_{z,y}(\mathbf{u}_2,\mathbf{u}_4) \\ C_{z,y}(\mathbf{u}_3,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_3,\mathbf{u}_2) & C_{y,y}(\mathbf{u}_3,\mathbf{u}_3) & C_y(\mathbf{u}_3,\mathbf{u}_4) \\ C_{z,y}(\mathbf{u}_4,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_4,\mathbf{u}_2) & C_{y,y}(\mathbf{u}_4,\mathbf{u}_3) & C_y(\mathbf{u}_4,\mathbf{u}_4) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} C_z(\mathbf{u}_o,\mathbf{u}_1) \\ C_z(\mathbf{u}_o,\mathbf{u}_2) \\ C_{z,y}(\mathbf{u}_o,\mathbf{u}_3) \\ C_{z,y}(\mathbf{u}_o,\mathbf{u}_4) \end{bmatrix}$$

$$\begin{bmatrix} C_z(\mathbf{u}_1,\mathbf{u}_1) & C_z(\mathbf{u}_1,\mathbf{u}_n) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_o) \\ C_z(\mathbf{u}_n,\mathbf{u}_1) & C_z(\mathbf{u}_n,\mathbf{u}_n) & C_{z,y}(\mathbf{u}_n,\mathbf{u}_o) \\ C_{z,y}(\mathbf{u}_0,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_0,\mathbf{u}_n) & C_y(\mathbf{u}_o,\mathbf{u}_o) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_n \\ \lambda_y \end{bmatrix} = \begin{bmatrix} C_z(\mathbf{u}_o,\mathbf{u}_1) \\ C_z(\mathbf{u}_o,\mathbf{u}_n) \\ C_{z,y}(\mathbf{u}_o,\mathbf{u}_o) \end{bmatrix}$$

Full cokriging system (above) and collocated cokriging system (below) to 2 primary data and only the collocated secondary data.

COLLOCATED COKRIGING

- The collocated value is surely the most important and likely screens the influence of multiple secondary data
  - Therefore, secondary variogram is not needed, never need covariance between secondary data

$$\begin{bmatrix} C_z(\mathbf{u}_1,\mathbf{u}_1) & C_z(\mathbf{u}_1,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_o) \\ C_z(\mathbf{u}_2,\mathbf{u}_1) & C_z(\mathbf{u}_2,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_2,\mathbf{u}_o) \\ C_{z,y}(\mathbf{u}_0,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_0,\mathbf{u}_2) & \boxed{\sigma_y^2} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_n \\ \lambda_y \end{bmatrix} = \begin{bmatrix} C_z(\mathbf{u}_o,\mathbf{u}_1) \\ C_z(\mathbf{u}_o,\mathbf{u}_2) \\ C_{z,y}(0\ ) \end{bmatrix}$$

  - No need for secondary variogram, $\gamma_y(\mathbf{h})$.
  - For the secondary data on the left-hand redundancy size, we only need the sill, $C_y(0) = \sigma_y^2$

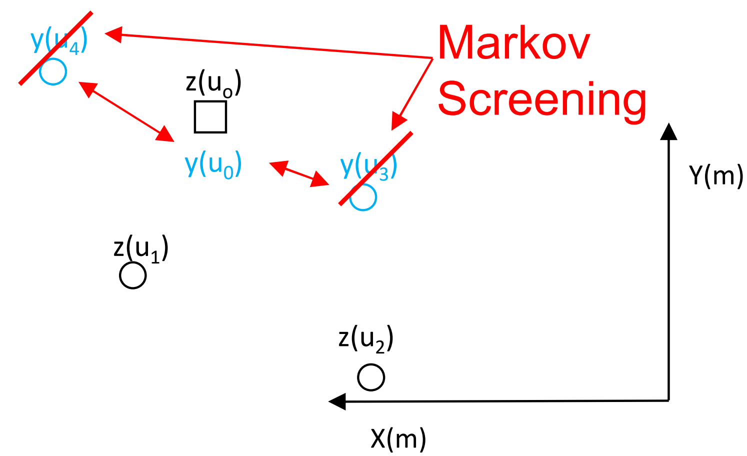


Illustration of the impact of Markov screening on the full cokriging system (above), non-collocated secondary data removed (below).

COLLOCATED COKRIGING

- The collocated value is surely the most important and likely screens the influence of multiple secondary data
  - Therefore, secondary variogram is not needed, never need covariance between secondary data

$$\begin{bmatrix} C_z(\mathbf{u}_1,\mathbf{u}_1) & C_z(\mathbf{u}_1,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_1,\mathbf{u}_o) \\ C_z(\mathbf{u}_2,\mathbf{u}_1) & C_z(\mathbf{u}_2,\mathbf{u}_2) & C_{z,y}(\mathbf{u}_2,\mathbf{u}_o) \\ C_{z,y}(\mathbf{u}_0,\mathbf{u}_1) & C_{z,y}(\mathbf{u}_0,\mathbf{u}_2) & \sigma_y^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_n \\ \lambda_y \end{bmatrix} = \begin{bmatrix} C_z(\mathbf{u}_o,\mathbf{u}_1) \\ C_z(\mathbf{u}_o,\mathbf{u}_2) \\ \boxed{\rho_{z,y}} \end{bmatrix}$$

  - No need for cross variogram, $\gamma_{z,y}(\mathbf{h})$ on closeness side, right-hand side of the kriging system.
  - Only need $C_{z,y}(0) = \rho_{z,y}$, the correlation coefficient.

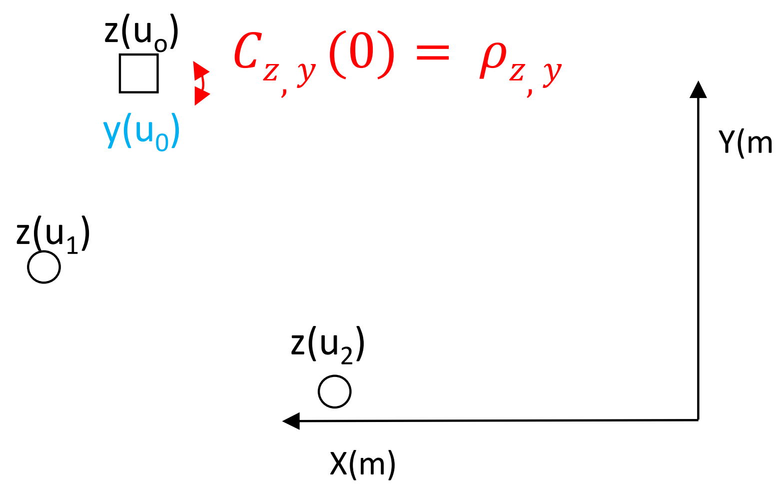
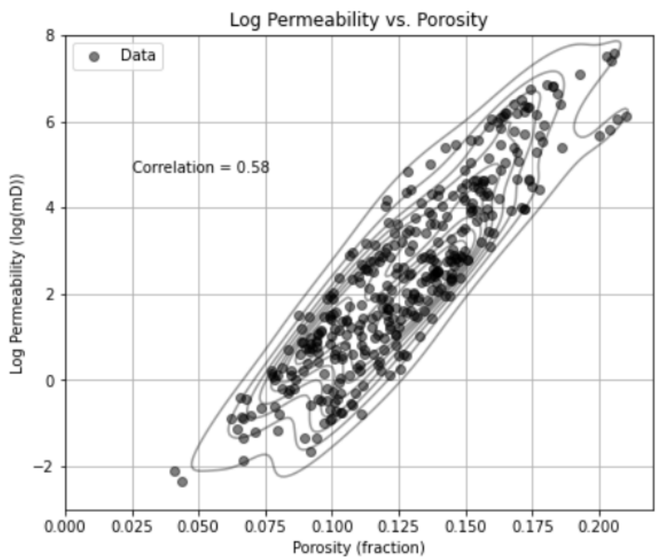
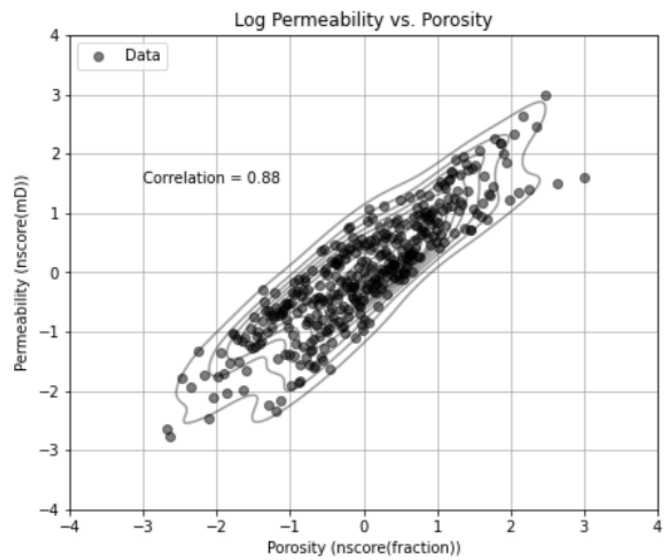


Illustration of the impact of Markov screening on the full cokriging system (above), non-collocated secondary data removed (below).

COLLOCATED COKRIGING

- The correlation coefficient for collocated cokriging.
  - kriging is applied in Gaussian space for the sequential Gaussian simulation paradigm.
  - therefore, we require the correlation coefficient in Gaussian space.
  - this is approximated by univariate transformations of both primary and secondary features.





Scatter plot and correlation coefficient for original (left) and Gaussian transformed features (right), non-collocated secondary data removed (below), file is GeostatsPy_cosimulation.ipynb.

  - often assists with outliers and improves the inference of the correlation coefficient
  - this is an approximation that could be, but is rarely ever, checked.

COLLOCATED COKRIGING

- Bayesian updating to calculate the cross variogram from the primary variogram x correlation coefficient between primary and secondary features.

$$\gamma_{z,y}(\mathbf{h}) = \rho_{z,y}\gamma_z(\mathbf{h}) \qquad C_{z,y}(\mathbf{h}) = \rho_{z,y} - \gamma_{z,y}(\mathbf{h})$$

  - Now the cross variogram is not needed on the left-hand side, redundancy of the kriging system either!

$$\begin{bmatrix} C_z(\mathbf{u}_1,\mathbf{u}_1) & C_z(\mathbf{u}_1,\mathbf{u}_2) & \rho_{z,y}C_z(\mathbf{u}_1,\mathbf{u}_o) \\ C_z(\mathbf{u}_2,\mathbf{u}_1) & C_z(\mathbf{u}_2,\mathbf{u}_2) & \rho_{z,y}C_z(\mathbf{u}_2,\mathbf{u}_o) \\ \rho_{z,y}C_z(\mathbf{u}_0,\mathbf{u}_1) & \rho_{z,y}C_z(\mathbf{u}_0,\mathbf{u}_2) & \sigma_y^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_n \\ \lambda_y \end{bmatrix} = \begin{bmatrix} C_z(\mathbf{u}_o,\mathbf{u}_1) \\ C_z(\mathbf{u}_o,\mathbf{u}_2) \\ \rho_{z,y} \end{bmatrix}$$

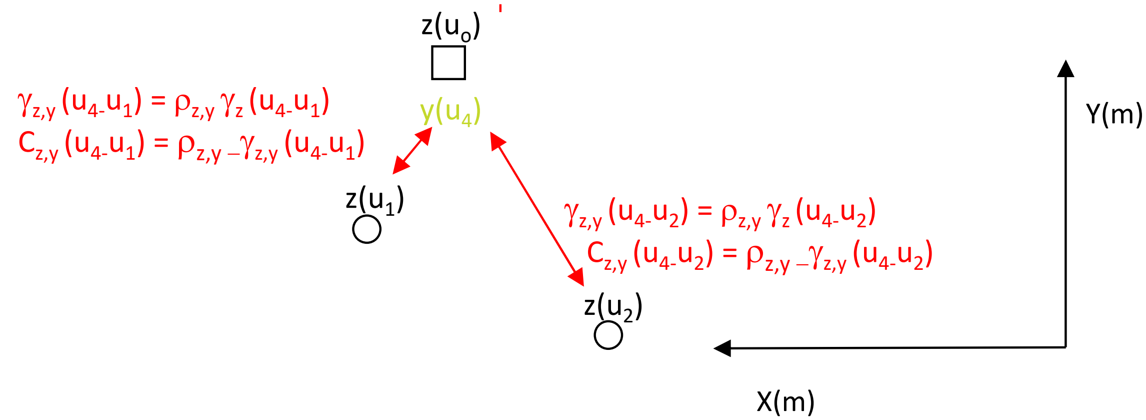


Illustration of the impact of Bayesian updating on the full cokriging system (above), and covariances between primary data and collocated secondary (below).

COLLOCATED COKRIGING

- Simplification due to retaining the collocated secondary variable:
  - no need for the variogram of the secondary variable
  - no need for a cross variogram
  - if the secondary data are smooth then considering more than the collocated variable should not help

- There is a **potential problem** with excess variance of the results when used in simulation mode
  - resulting mean is often biased if variance is too high
  - practical approach is to set a constant variance reduction factor (a global correction over all estimation locations)

COLLOCATED COKRIGING

- Eliminate the need for calculating a cross variogram:
  - no need for a cross variogram, i.e., the linear model of coregionalization

- We are [Bayesian] updating the primary variogram with the primary to secondary correlation coefficient as the new sill.

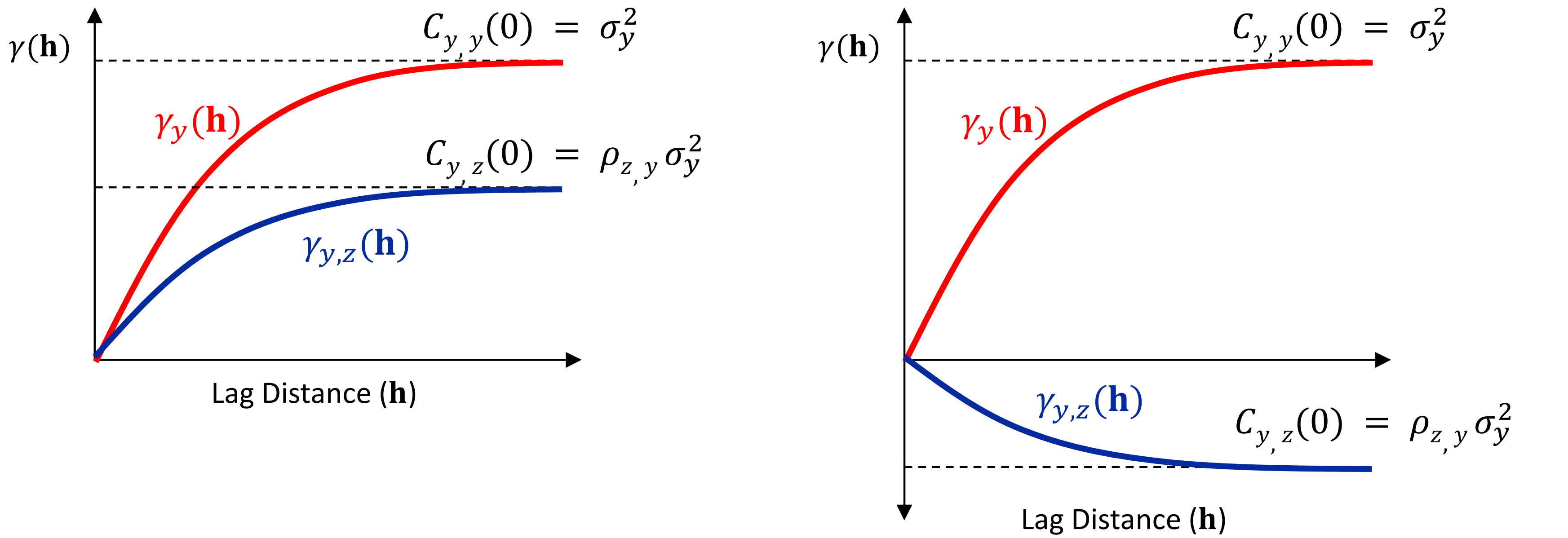
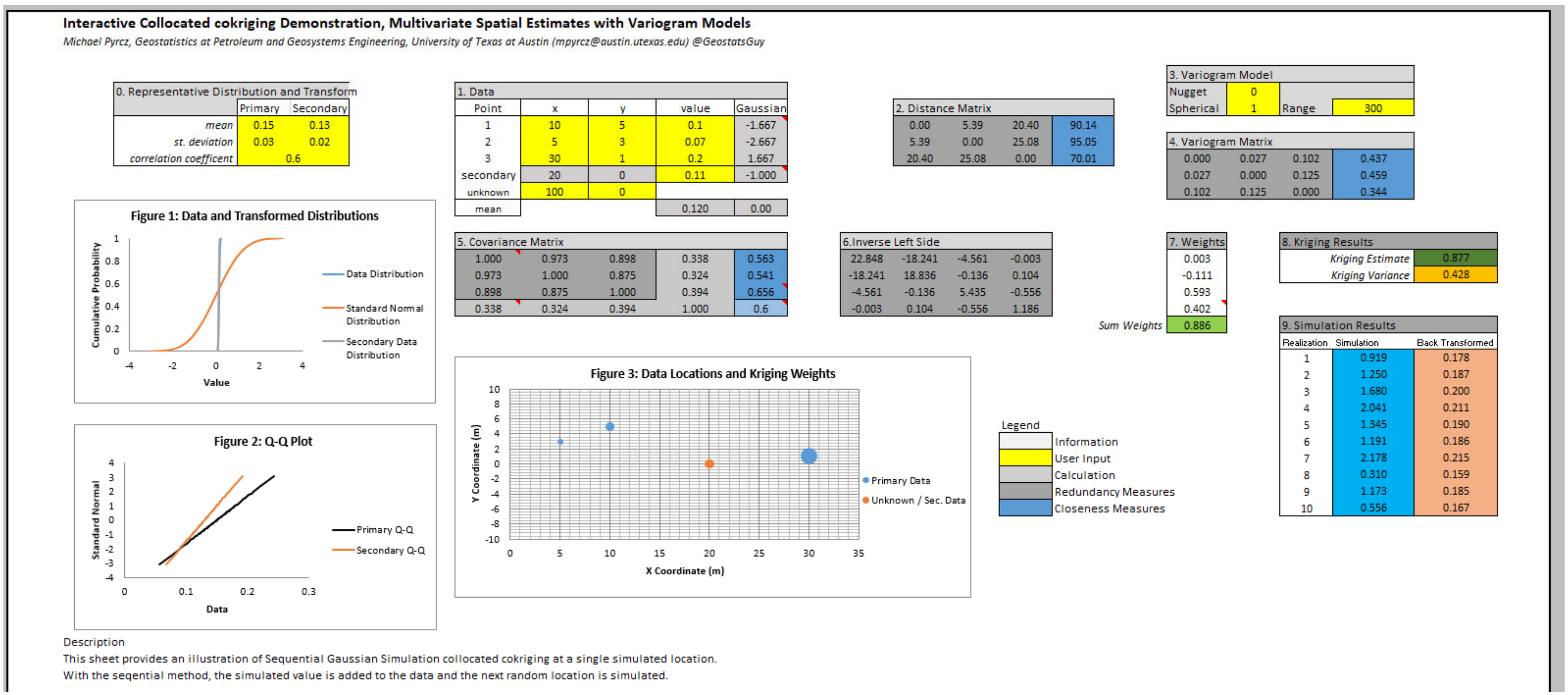


Illustration Bayesian updating the primary variogram to approximate the cross variogram for positive (left) and negative primary to secondary correlation coefficients.

COLLOCATED COKRIGING HANDS ON

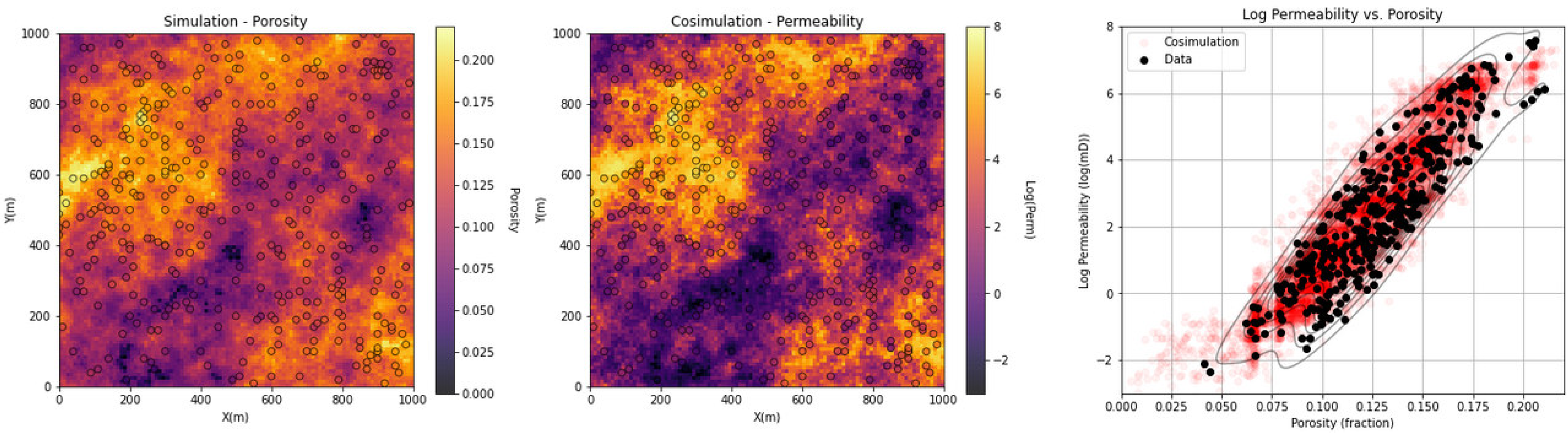
- Hands-on example of collocated cokriging:



Kriging and simulation at a single location with collocated cokriging in Excel, file is Collocated_Cokriging_Demo.xlsx

COLLOCATED COKRIGING HANDS ON

- Co-simulation Workflow in Python:
- Walkthrough and try to:
  - Change the correlation coefficient for co-simulation.
  - Try changing the variogram range and check the scatter plot for independent simulation



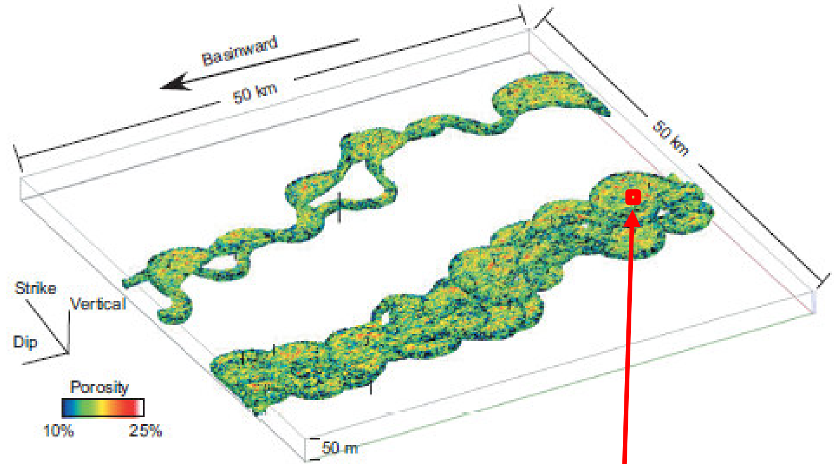
Cosimulation of permeability (center) given porosity (left) and cross plot data and simulated values (right), file is GeostatsPy_cosimulation.ipynb.

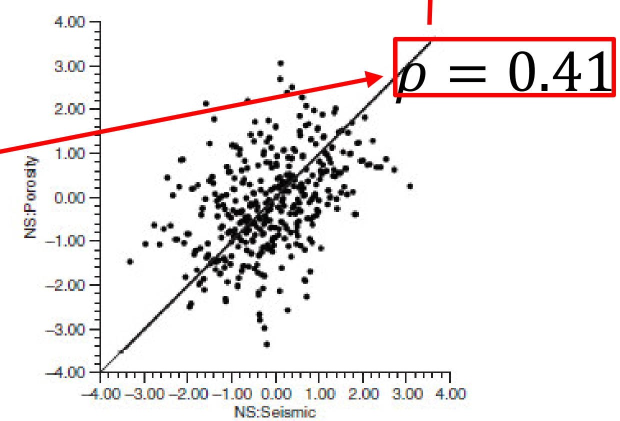
COMMON MODELING WORKFLOW

- The Co-simulation Workflow:

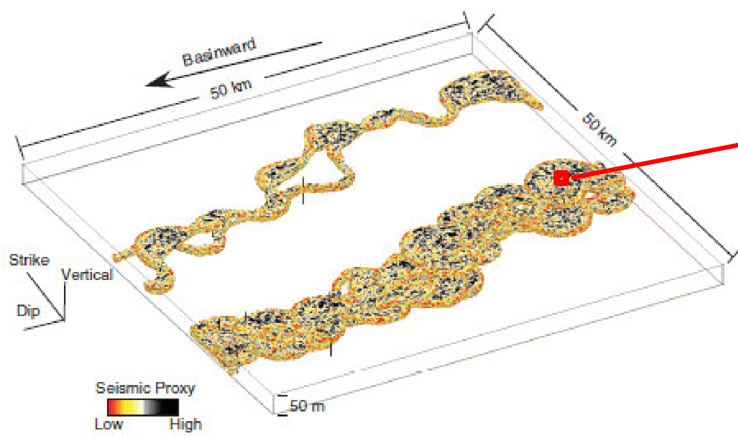
1. Simulate realization of secondary variable (e.g., acoustic impedance)
2. Integrate the collocated secondary realization at each location with correlation coefficient.
3. Check simulation histogram, variogram and scatter plot.

3. Exhaustive Primary Cosimulation





2. Correlation coefficient after Gaussian transform of primary and secondary variable



1. Exhaustive Secondary Simulation

Cosimulation of porosity given a realization of acoustic impedance and Gaussian space correlation coefficient.

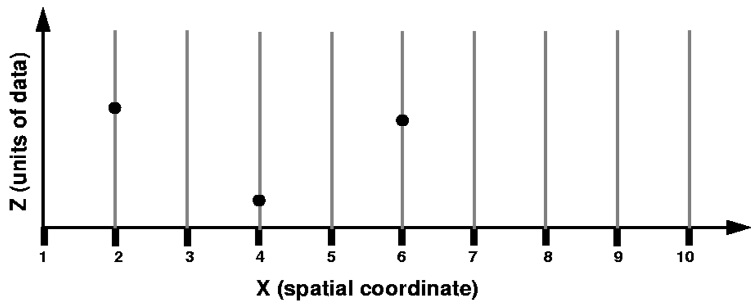
P-FIELD SIMULATION

P-FIELD AND CLOUD TRANSFORM

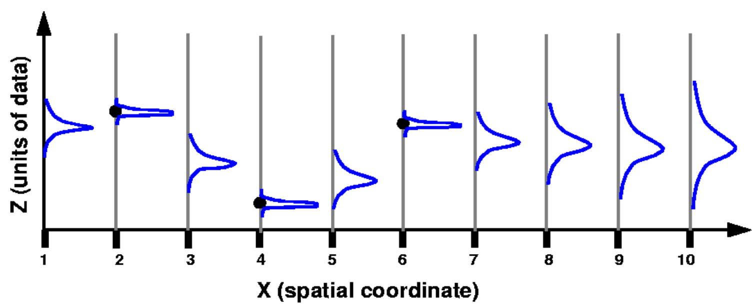
- The key idea of p-field simulation is to perform the simulation in two separate steps:
    - Construct local distributions of uncertainty
    - Draw from those distributions simultaneously with correlated probabilities
- Separating the two steps simulation has advantages:
    - the distributions of uncertainty can be constructed to honor all data and checked before any realizations are drawn, and
    - the simulations are consistent with the distributions of uncertainty
- Also, good reproduction of the primary to secondary data scatter plot.
- Some disadvantages include potentially poor reproduction of the histogram and variogram.
- The most commonly used approach is known as cloud transform, often used to simulate permeability from porosity

P-FIELD SIMULATION

(a) Data values



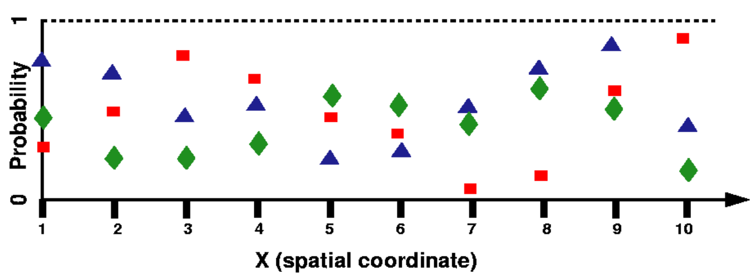
(b) Distributions of uncertainty



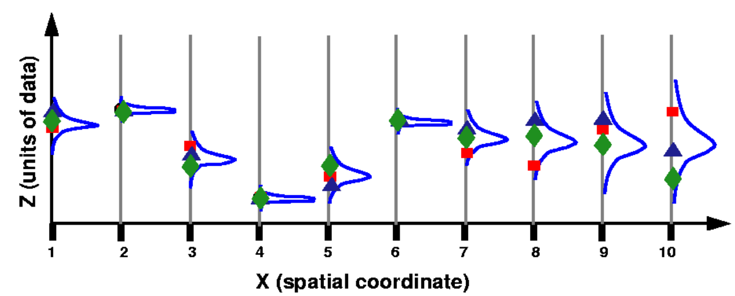
- The p-field simulation method steps:

1. Construct local distributions of uncertainty with all available information sources
2. Generate spatially correlated probability values (p-values) (usual with Gaussian simulation) uniform [0,1] distributed to avoid bias.
3. Draw values simultaneously and retain together as a realization

(c) Probability values (unconditional simulations)



(d)Simulated values drawn from conditional distributions

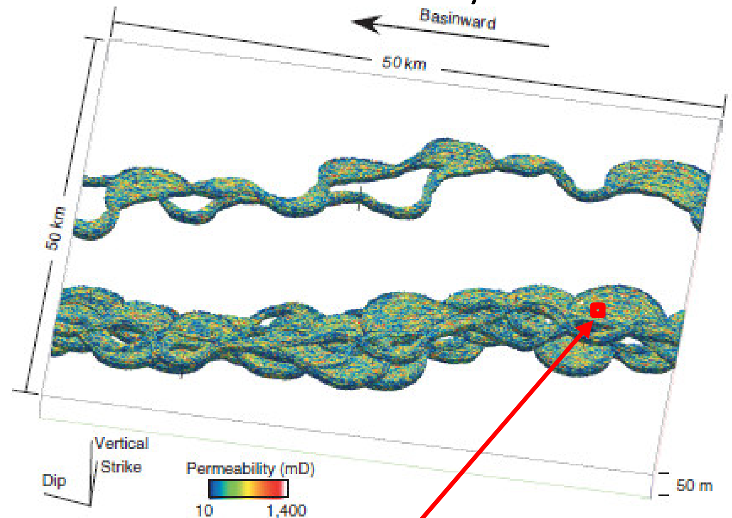


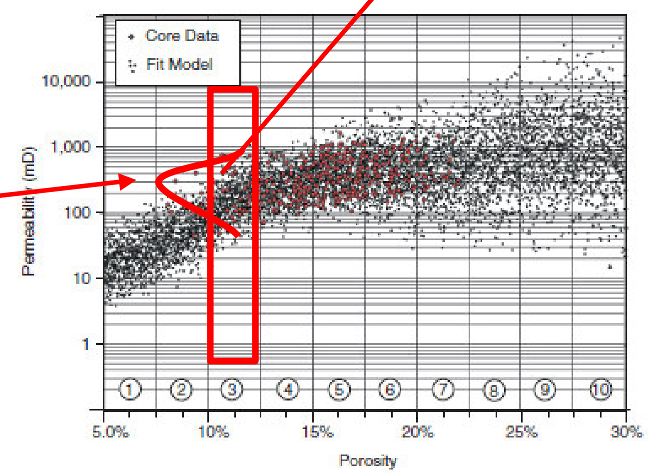
CLOUD TRANSFORM

- Cloud Transform Workflow

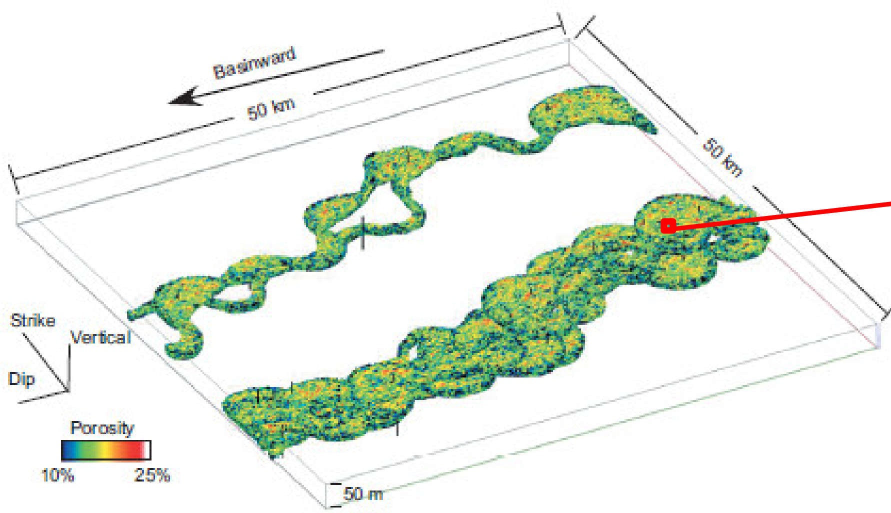
1. Simulate realization of secondary variable (e.g., porosity)
2. For each location draw from the conditional distribution given the secondary realization with spatially correlated p-values
3. Check model histogram, variogram and bivariate plot.

3. Exhaustive Primary Cosimulation





2. Conditional Primary | Secondary



1. Exhaustive Secondary Simulation

Cloud transform simulation of permeability given a realization of porosity and permeability | porosity conditional distributions.

CO-SIMULATION SUMMARY

- What should you have learned?
  - Co-simulation includes methods that simulate a property realization (primary) conditional to a previously simulated property realization (secondary)
  - Two independently simulated features [without co-simulation] will only have correlations imposed by data and outside spatial correlation, variogram range of data will be uncorrelated.
  - This leads to implausible combinations of the features and inflated spatial uncertainty.

- Two methods are commonly applied for co-simulation:
  1. Collocated Cokriging simplifies the full Cokriging system
     - Markov assumption – only need collocated secondary value
     - Bayesian updating – get the cross variogram by scaling the primary variogram with correlation coefficient
     - May not reproduce the cloud well (relationship between the 2 variables)
  2. Cloud transform
     - Forces reproduction of the cloud
     - May not get the spatial continuity and distribution well

- It is essential to know their assumptions and steps

CO-SIMULATION

New Tools

Topic	Application to Subsurface Modeling
Full Cokriging	Use variograms, cross variograms to model multivariate spatial phenomenon. *We could do it right!*
Collocated Cokriging	Use variogram and correlation coefficient with the Markov / Bayes assumptions. A practical, commonly used workflow for bivariate co-simulation.

Co-simulation

Lecture outline . . .

- Co-simulation
- Full Cokriging
- Collocated Cokriging
- P-field Simulation