Open Source Spatial Data Analytics in Python with GeostatsPy

Introduction to GeostatsPy

Lecture outline . . .

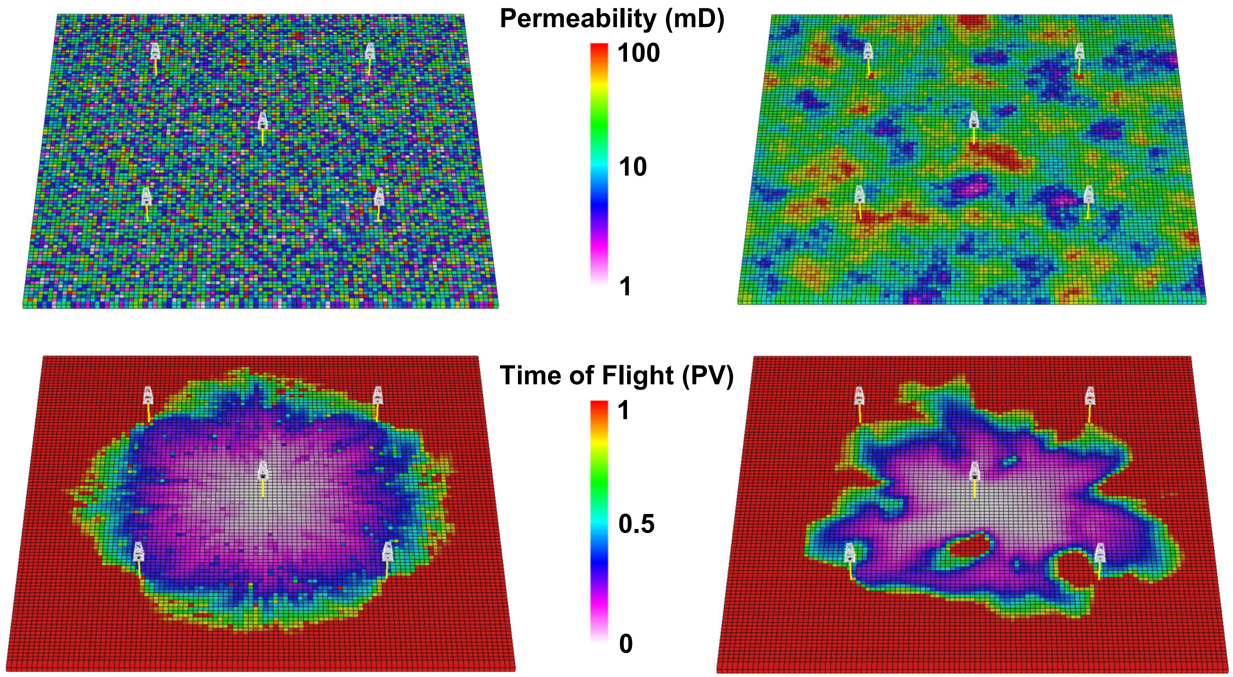
- **Variogram Calculation**
- **Interactive Demo with GeostatsPy**
- **Workflow with GeostatsPy**

Michael Pyrcz, The University of Texas at Austin

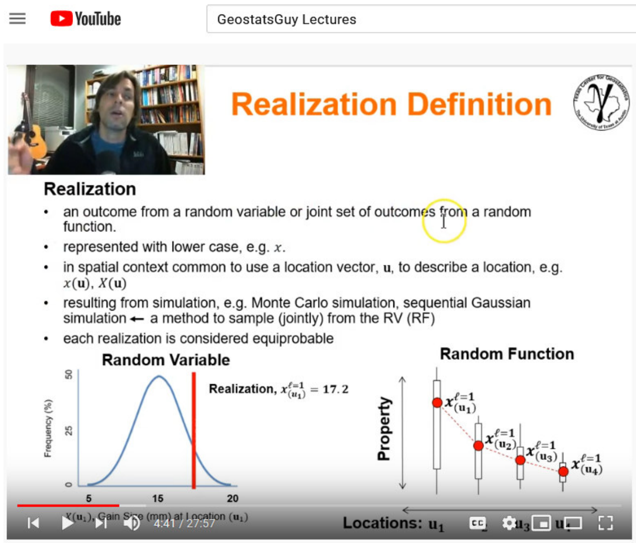
Motivation

Spatial continuity has a significant impact on reservoir flow rates and overall recovery.

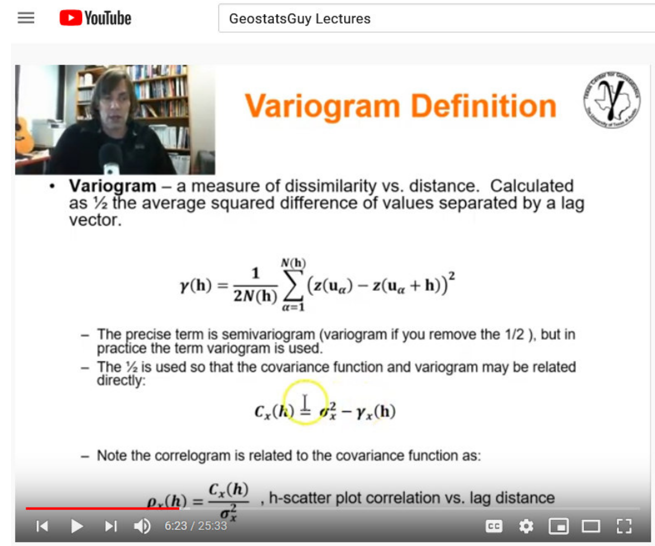
- We need to quantify spatial continuity and predict / forecast with it.



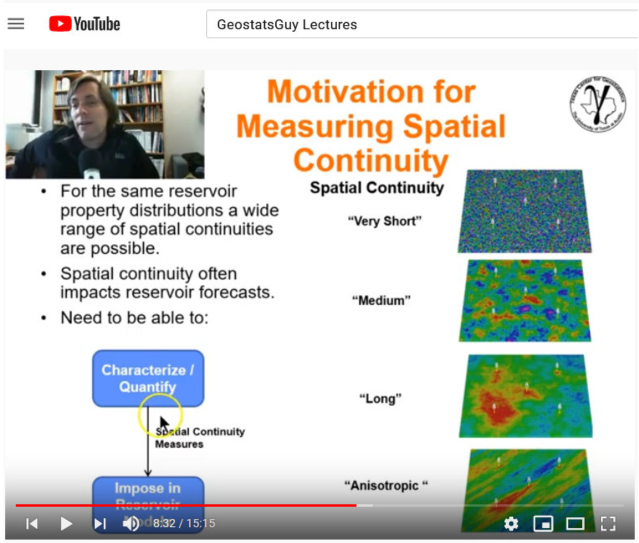
Recorded Lectures



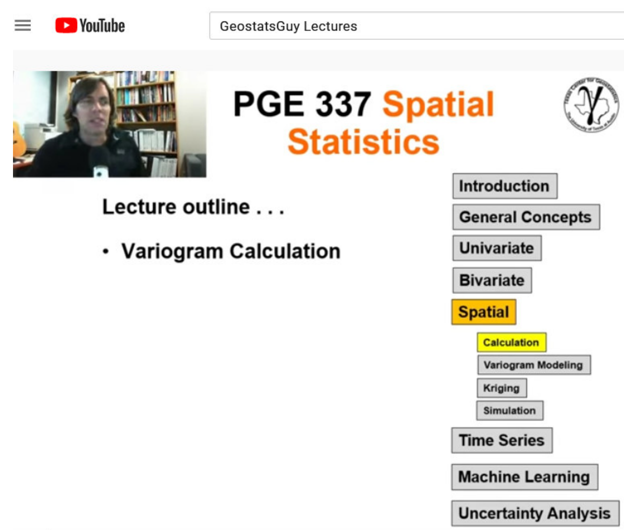
10 Data Analytics: Spatiotemporal Stationarity



10c Data Analytics: Variogram Introduction



10b Data Analytics: Spatial Continuity



10d Data Analytics: Variogram Calculation

Michael Pyrcz, The University of Texas at Austin

Introduction to GeostatsPy

Lecture outline . . .

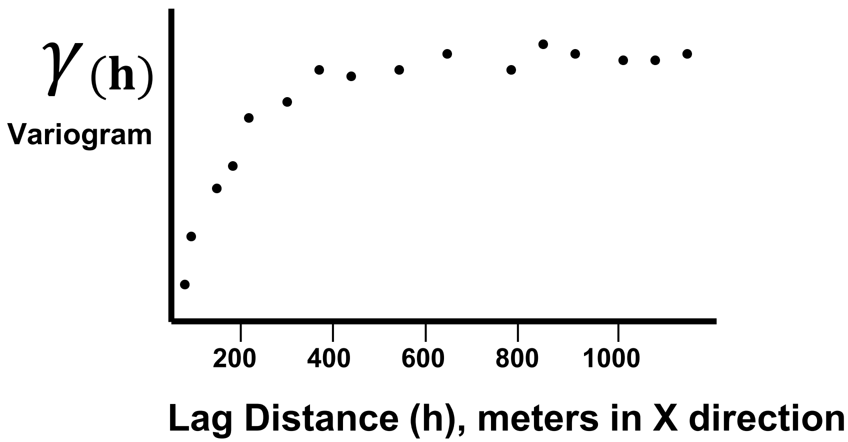
- **Variogram Calculation**

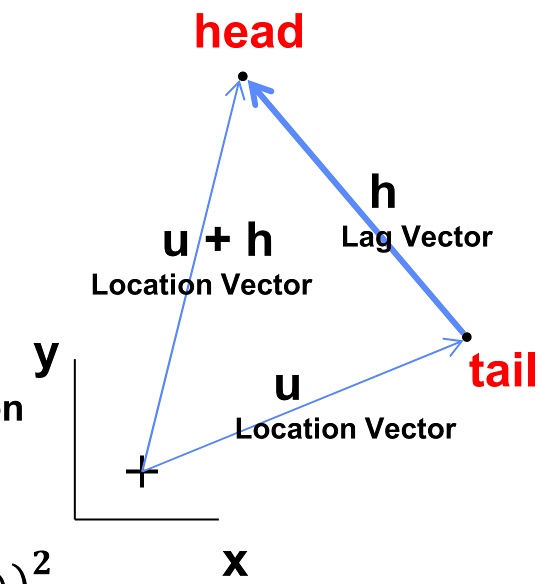
Measuring Spatial Continuity

We need a statistic to quantify spatial continuity!

The Semivariogram:

- Function of difference over distance.





- The equation:

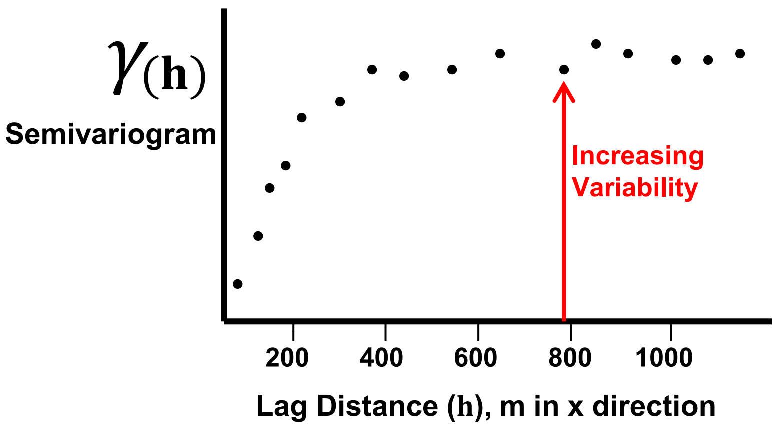
$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} \left(\overset{\text{tail}}{z(\mathbf{u}_\alpha)} - \overset{\text{head}}{z(\mathbf{u}_\alpha + \mathbf{h})} \right)^2$$

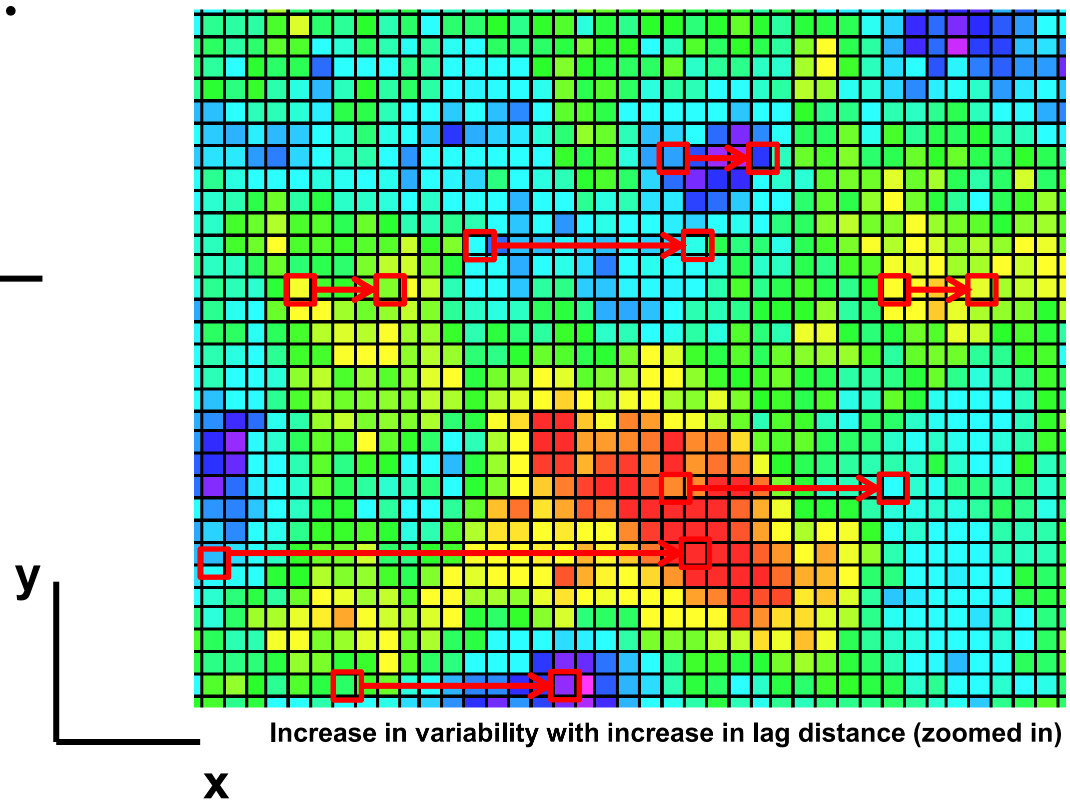
One half the average squared difference over lag distance, $\mathbf{h}$, over all possible pairs of data, $N(\mathbf{h})$.

Variogram Observations

Observation #1

- As distance increases, variability increase (in general).





Increase in variability with increase in lag distance (zoomed in)

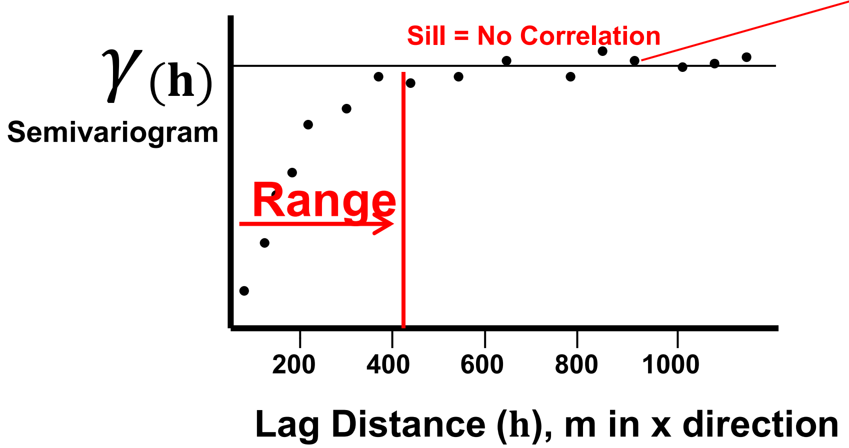
Michael Pyrcz, The University of Texas at Austin

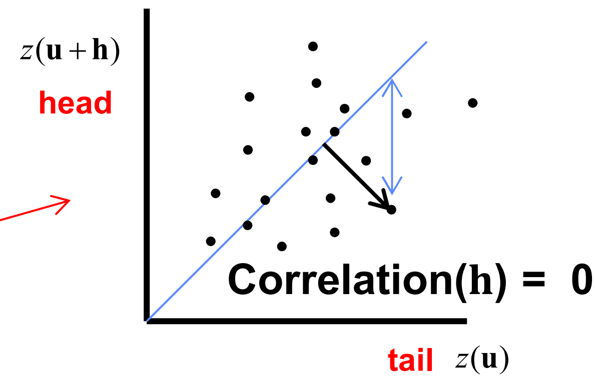
Variogram Observations

Observation #4

- The lag distance at which the variogram reaches the sill is know as the range.

  - At the range, knowing the data value at the tail provides no information about a value at the head.





Michael Pyrcz, The University of Texas at Austin

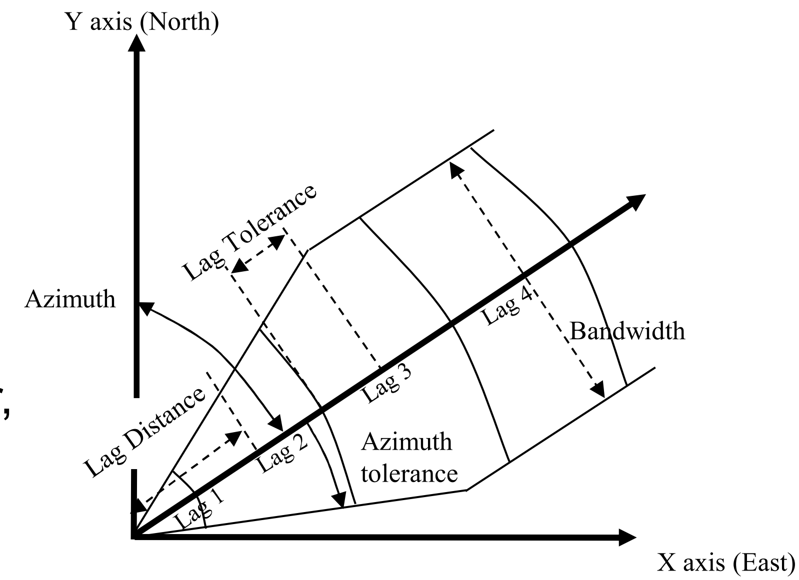
Calculating Experimental Variograms

How do we get pairs separated by lag vector?

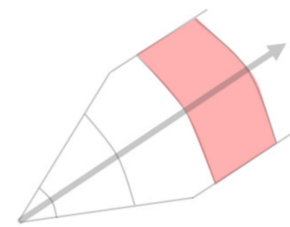
We need to find pairs with the same lag vector, same distance and direction

- Irregular spaced data in 2D:
  - Unit lag distance and distance tolerance
  - Azimuth direction and azimuth tolerance
  - Bandwidth (maximum deviation) from azimuth

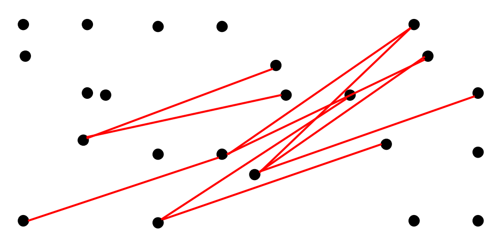
- We will calculate (and then model) the variogram in major and minor directions.



Variogram search template from Pyrcz and Deutsch, 2014



4th lag, 4 x lag distance in 045 azimuth



All pairs for 4th lag 045 azimuth.

Michael Pyrcz, The University of Texas at Austin

Introduction to GeostatsPy

Lecture outline . . .

- **Interactive Demo with GeostatsPy**

Interactive Variogram Calculation Demonstration with GeostatsPy

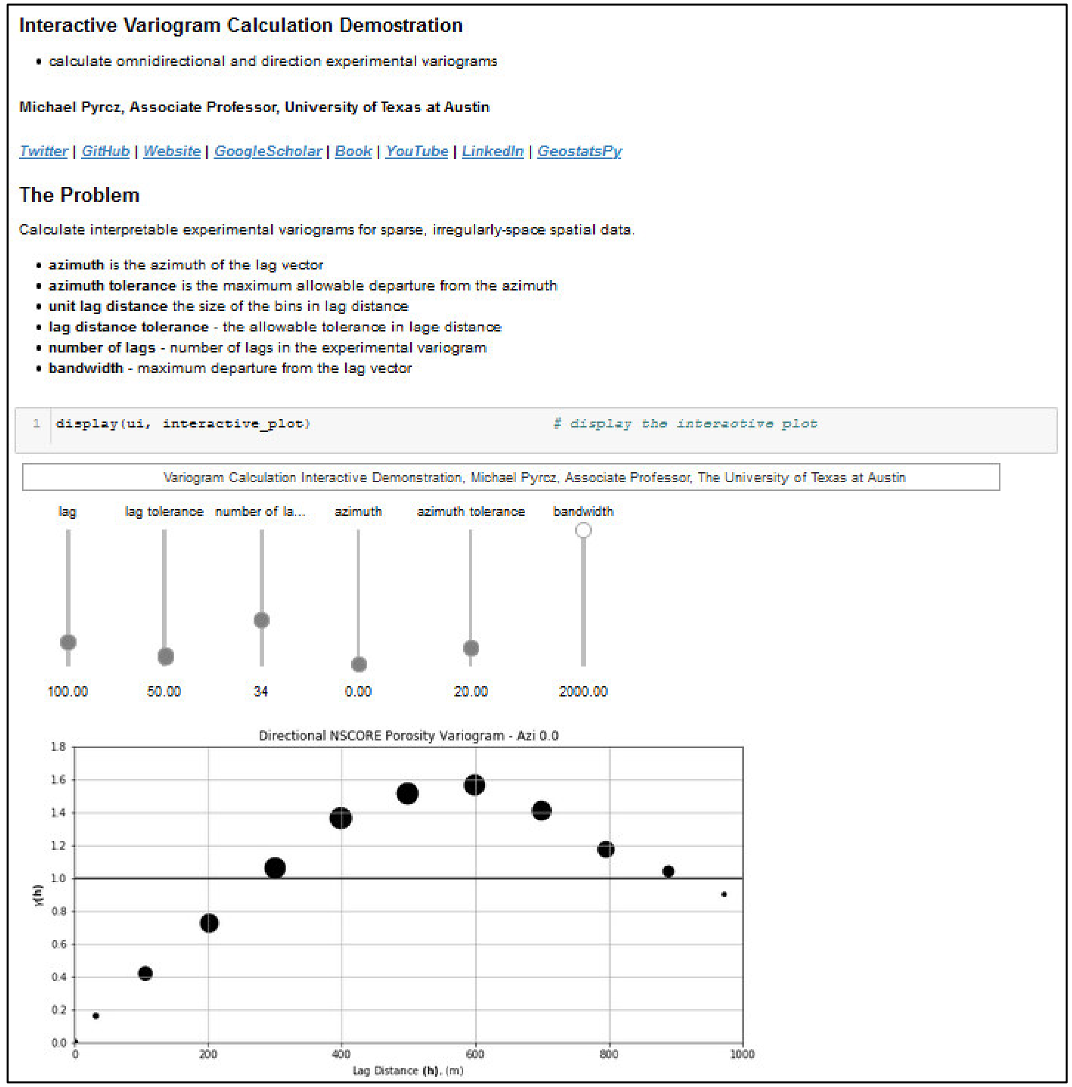
Let's calculate variograms:

- normal score transformed (standard Gaussian) porosity

Some Hints:

- start with lag of the minimum data spacing
- start with lag tolerance as ½ lag
- start isotopic with azimuth tolerance of 90

Find and describe the major and minor direction variograms.



Interactive Python Jupyter variogram calculation (Interactive_Variogram_Calculation.ipynb).

Michael Pyrcz, The University of Texas at Austin

Introduction to GeostatsPy

Lecture outline . . .

- **Workflow with GeostatsPy**

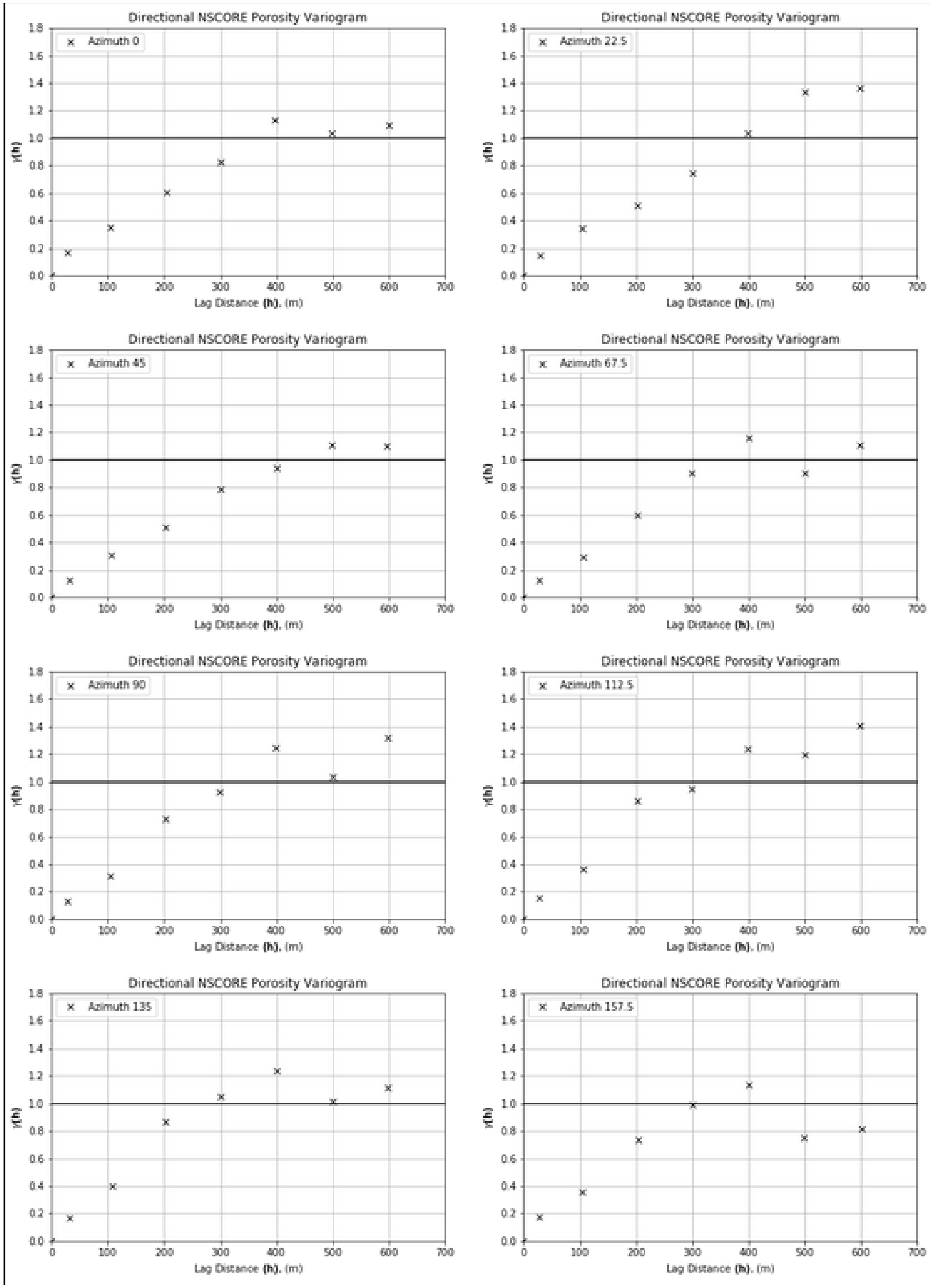
Michael Pyrcz, The University of Texas at Austin

Variogram Calculation Workflow with GeostatsPy

Let's walkthrough a more thorough variogram-based spatial analysis workflow:

- identifying the major and minor directions with GeostatsPy
- directional variograms and variogram maps.

Python Jupyter variogram calculation (GeostatsPy_spatial_continuity_directions.ipynb).



Michael Pyrcz, The University of Texas at Austin

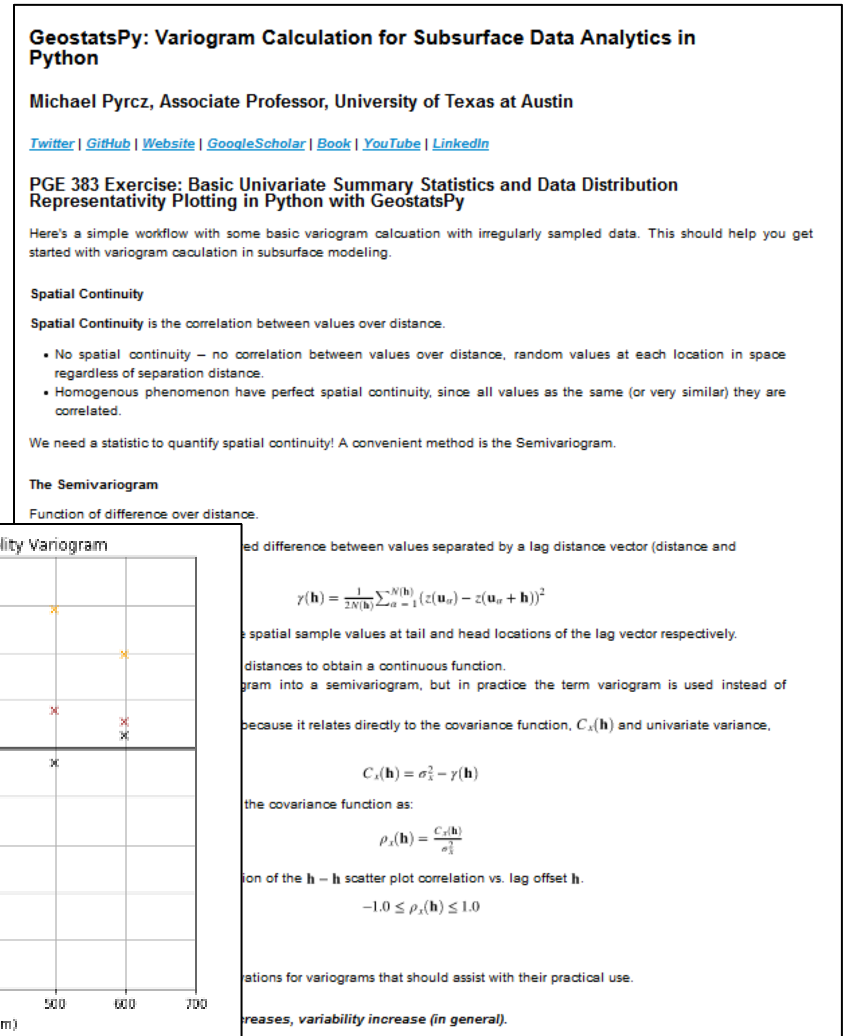
Facies-based Variograms Workflow with GeostatsPy

Variogram Calculation Workflow in Python

Walkthrough and try to:

- Change the tolerances and to calculate anisotropic experimental variograms.

- File is: GeostatsPy_variogram_calculation.ipynb



GeostatsPy: Variogram Calculation for Subsurface Data Analytics in Python

Michael Pyrcz, Associate Professor, University of Texas at Austin

Twitter | GitHub | Website | GoogleScholar | Book | YouTube | LinkedIn

PGE 383 Exercise: Basic Univariate Summary Statistics and Data Distribution Representativity Plotting in Python with GeostatsPy

Here's a simple workflow with some basic variogram calcuation with irregularly sampled data. This should help you get started with variogram caculation in subsurface modeling.

Spatial Continuity

Spatial Continuity is the correlation between values over distance.

- No spatial continuity – no correlation between values over distance, random values at each location in space regardless of separation distance.
- Homogenous phenomenon have perfect spatial continuity, since all values as the same (or very similar) they are correlated.

We need a statistic to quantify spatial continuity! A convenient method is the Semivariogram.

The Semivariogram

Function of difference over distance.

...ed difference between values separated by a lag distance vector (distance and

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})}\sum_{\alpha=1}^{N(\mathbf{h})}(z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h}))^2$$

...spatial sample values at tail and head locations of the lag vector respectively.

...distances to obtain a continuous function.

...gram into a semivariogram, but in practice the term variogram is used instead of

...because it relates directly to the covariance function, $C_x(\mathbf{h})$ and univariate variance,

$$C_x(\mathbf{h}) = \sigma_x^2 - \gamma(\mathbf{h})$$

...the covariance function as:

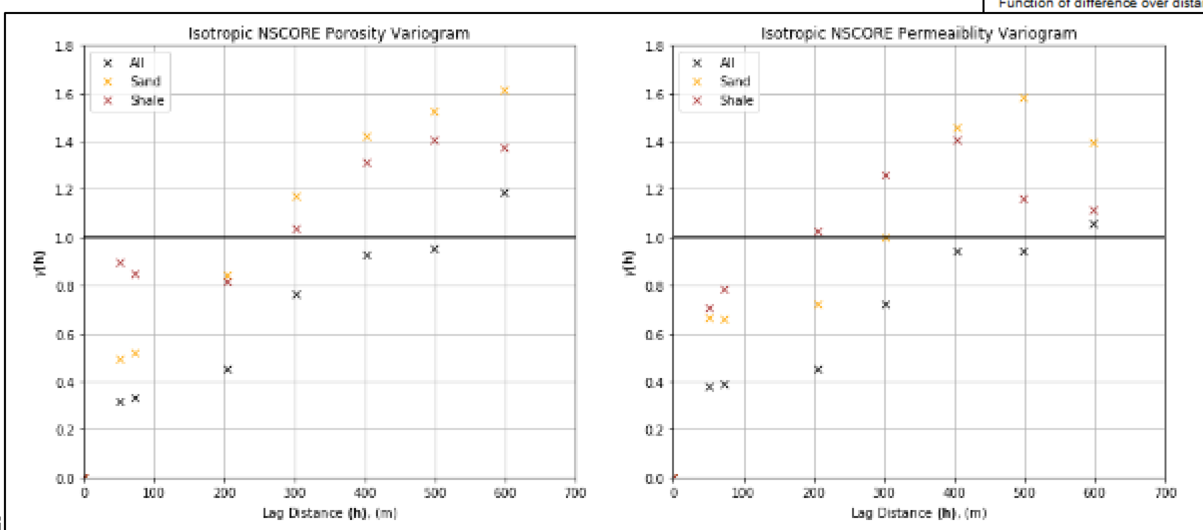
$$\rho_x(\mathbf{h}) = \frac{C_x(\mathbf{h})}{\sigma_x^2}$$

...ion of the $\mathbf{h}-\mathbf{h}$ scatter plot correlation vs. lag offset $\mathbf{h}$.

$$-1.0 \le \rho_x(\mathbf{h}) \le 1.0$$

...ations for variograms that should assist with their practical use.

...reases, variability increase (in general).



Open Source Spatial Data Analytics in Python with GeostatsPy

Introduction to GeostatsPy

Lecture outline . . .

- **Variogram Calculation**
- **Interactive Demo with GeostatsPy**
- **Workflow with GeostatsPy**

Michael Pyrcz, The University of Texas at Austin