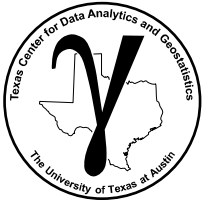


# Open Source Spatial Data Analytics in Python with GeostatsPy

## Introduction to GeostatsPy

Lecture outline . . .

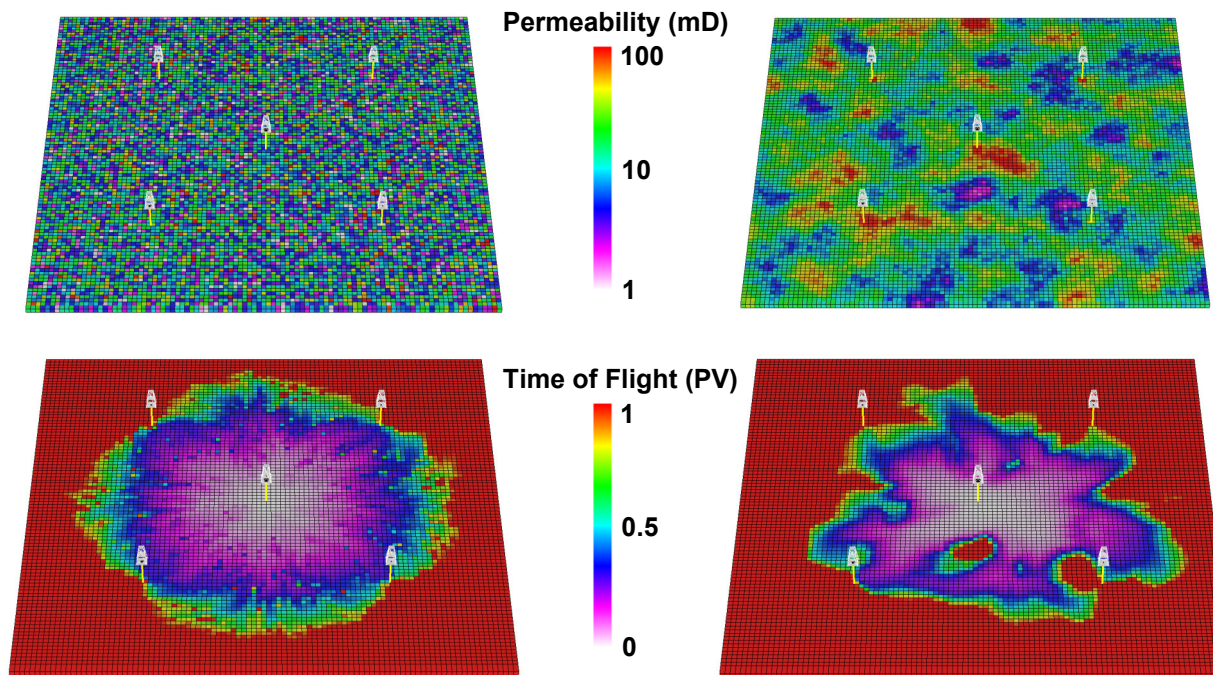
- Variogram Calculation
- Interactive Demo with GeostatsPy
- Workflow with GeostatsPy

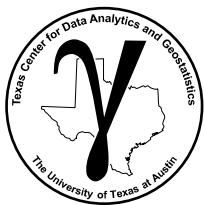


# Motivation

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

GeostatsGuy Lectures

## Realization Definition

**Realization**

- an outcome from a random variable or joint set of outcomes from a random function.
- represented with lower case, e.g.  $x$ .
- in spatial context common to use a location vector,  $u$ , to describe a location, e.g.  $x(u)$ ,  $X(u)$
- resulting from simulation, e.g. Monte Carlo simulation, sequential Gaussian simulation — a method to sample (jointly) from the RV (RF)
- each realization is considered equiprobable

**Random Variable**

Realization,  $x(u_1) = 17.2$

**Random Function**

10 Data Analytics: Spatiotemporal Stationarity

GeostatsGuy Lectures

## Motivation for Measuring Spatial Continuity

- For the same reservoir property distributions a wide range of spatial continuities are possible.
- Spatial continuity often impacts reservoir forecasts.
- Need to be able to:

Characterize / Quantify

Spatial Continuity Measures

Impose in Reservoir

**Spatial Continuity**

- "Very Short"
- "Medium"
- "Long"
- "Anisotropic"

10b Data Analytics: Spatial Continuity

GeostatsGuy Lectures

## Variogram Definition

- Variogram** — a measure of dissimilarity vs. distance. Calculated as  $\frac{1}{2}$  the average squared difference of values separated by a lag vector.

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

- The precise term is semivariogram (variogram if you remove the  $1/2$ ), but in practice the term variogram is used.
- The  $\frac{1}{2}$  is used so that the covariance function and variogram may be related directly:

$$C_x(h) = \sigma_x^2 - \gamma_x(h)$$

- Note the correlogram is related to the covariance function as:

$$\rho_x(h) = \frac{C_x(h)}{\sigma_x^2}, \text{ h-scatter plot correlation vs. lag distance}$$

10c Data Analytics: Variogram Introduction

GeostatsGuy Lectures

## PGE 337 Spatial Statistics

**Lecture outline . . .**

- Variogram Calculation

**Introduction**

**General Concepts**

**Univariate**

**Bivariate**

**Spatial**

**Calculation**

Variogram Modeling

Kriging

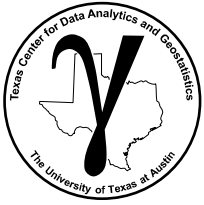
Simulation

**Time Series**

**Machine Learning**

**Uncertainty Analysis**

10d Data Analytics: Variogram Calculation

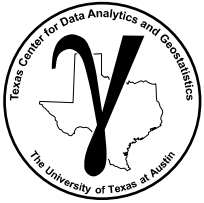


# Open Source Spatial Data Analytics in Python with GeostatsPy

## Introduction to GeostatsPy

Lecture outline . . .

- Variogram Calculation

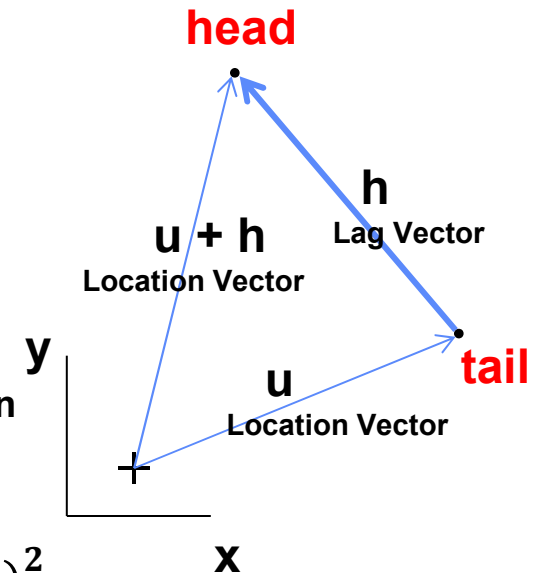
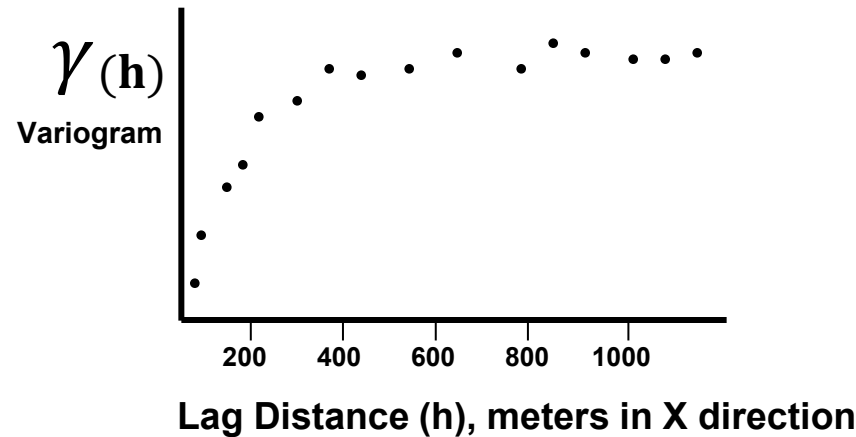


# Measuring Spatial Continuity

We need a statistic to quantify spatial continuity!

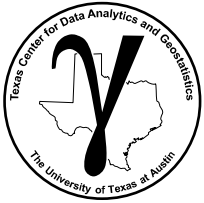
## The Semivariogram:

- Function of difference over distance.



- The equation: 
$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} (z(u_{\alpha}) - z(u_{\alpha} + h))^2$$

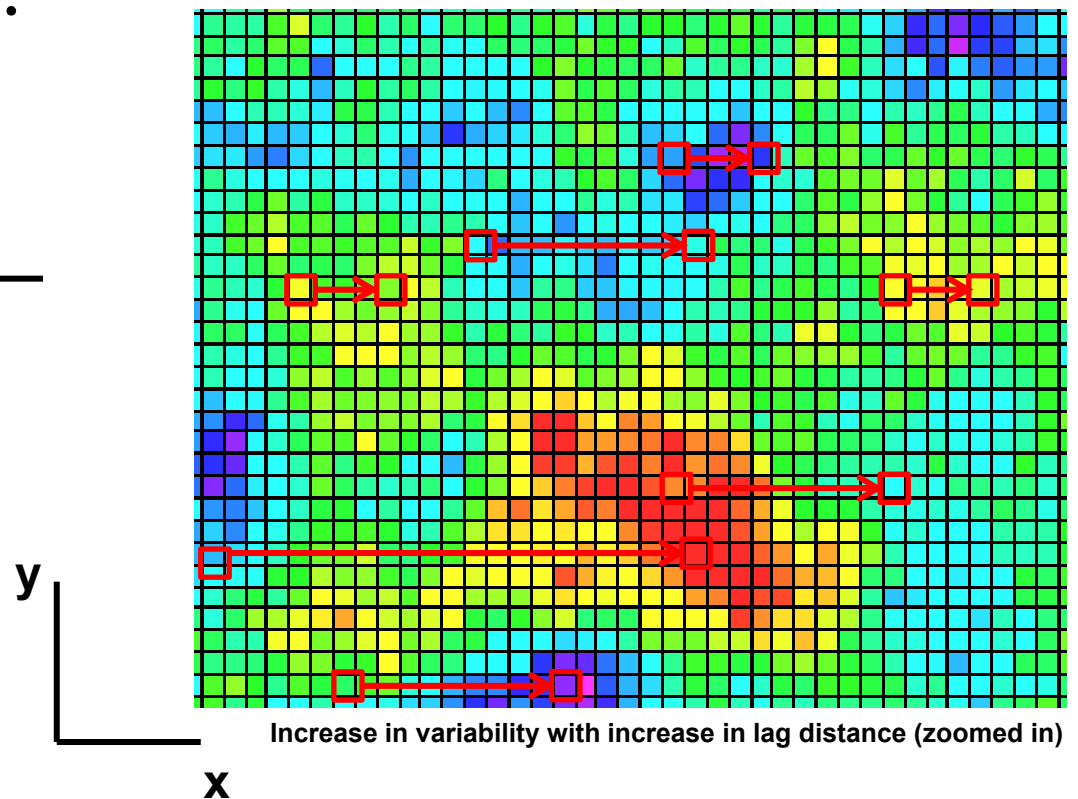
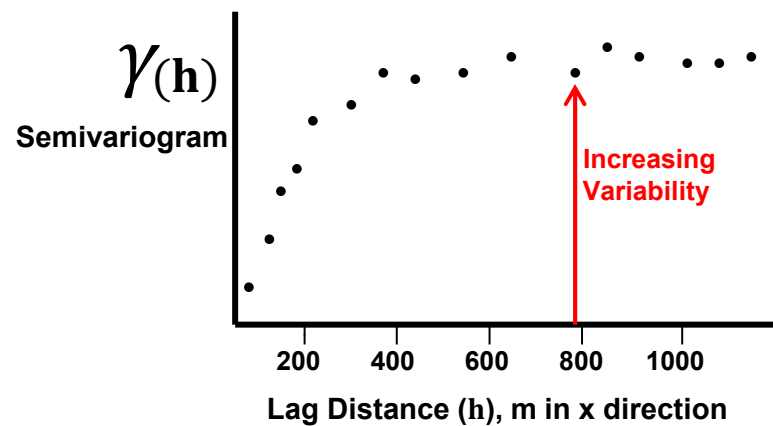
One half the average squared difference over lag distance,  $h$ , over all possible pairs of data,  $N(h)$ .

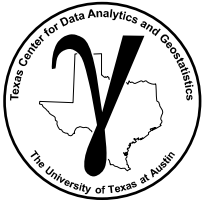


# Variogram Observations

## Observation #1

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

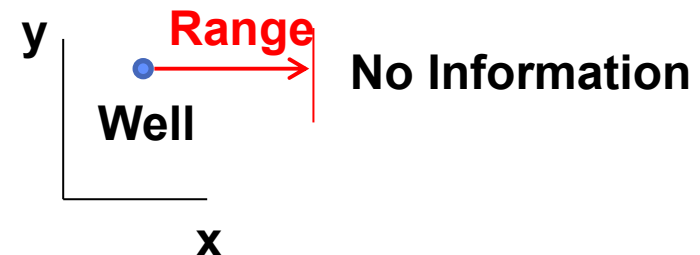
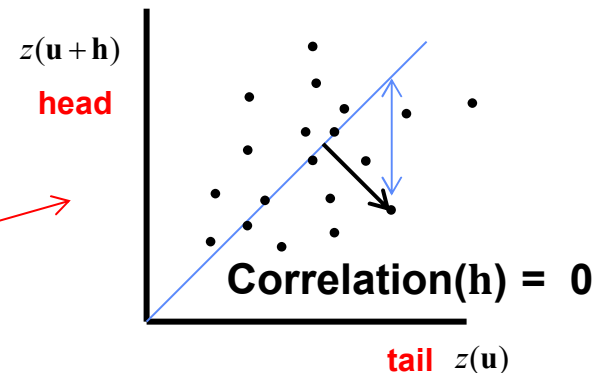
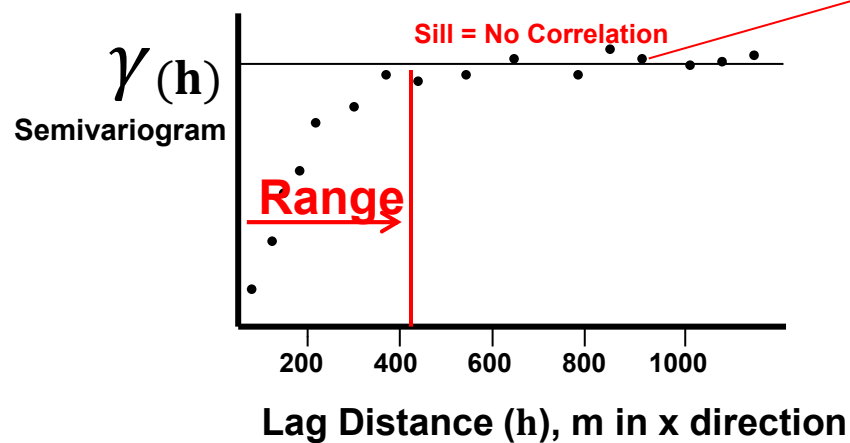


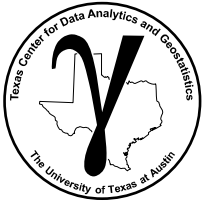


# Variogram Observations

## Observation #4

- The lag distance at which the variogram reaches the sill is known as the range.
- At the range, knowing the data value at the tail provides no information about a value at the head.



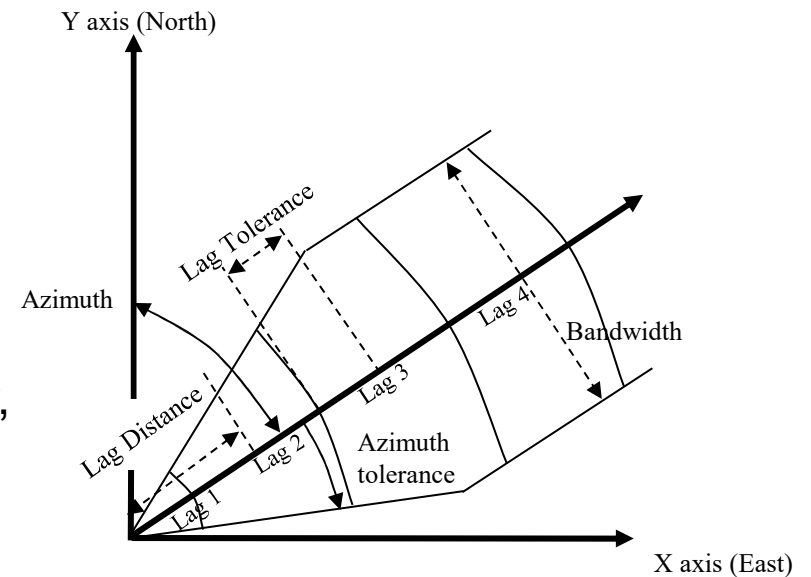


# Calculating Experimental Variograms

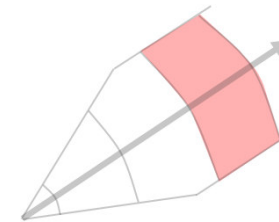
## How do we get pairs separated by lag vector?

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

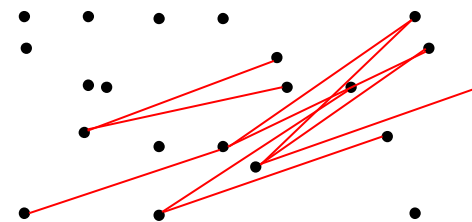
- Irregular spaced data in 2D:
  - Unit lag distance and distance tolerance
  - Azimuth direction and azimuth tolerance
  - Bandwidth (maximum deviation) from azimuth
- We will calculate (and then model) the variogram in major and minor directions.



Variogram search template from Pyrcz and Deutsch, 2014

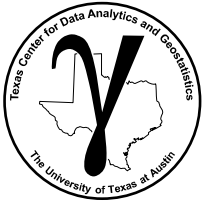


4<sup>th</sup> lag, 4 x lag distance in 045 azimuth



All pairs for 4<sup>th</sup> lag 045 azimuth.



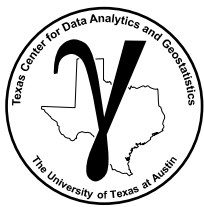


# Open Source Spatial Data Analytics in Python with GeostatsPy

## Introduction to GeostatsPy

Lecture outline . . .

- Interactive Demo with GeostatsPy



# Interactive Variogram Calculation Demonstration with GeostatsPy

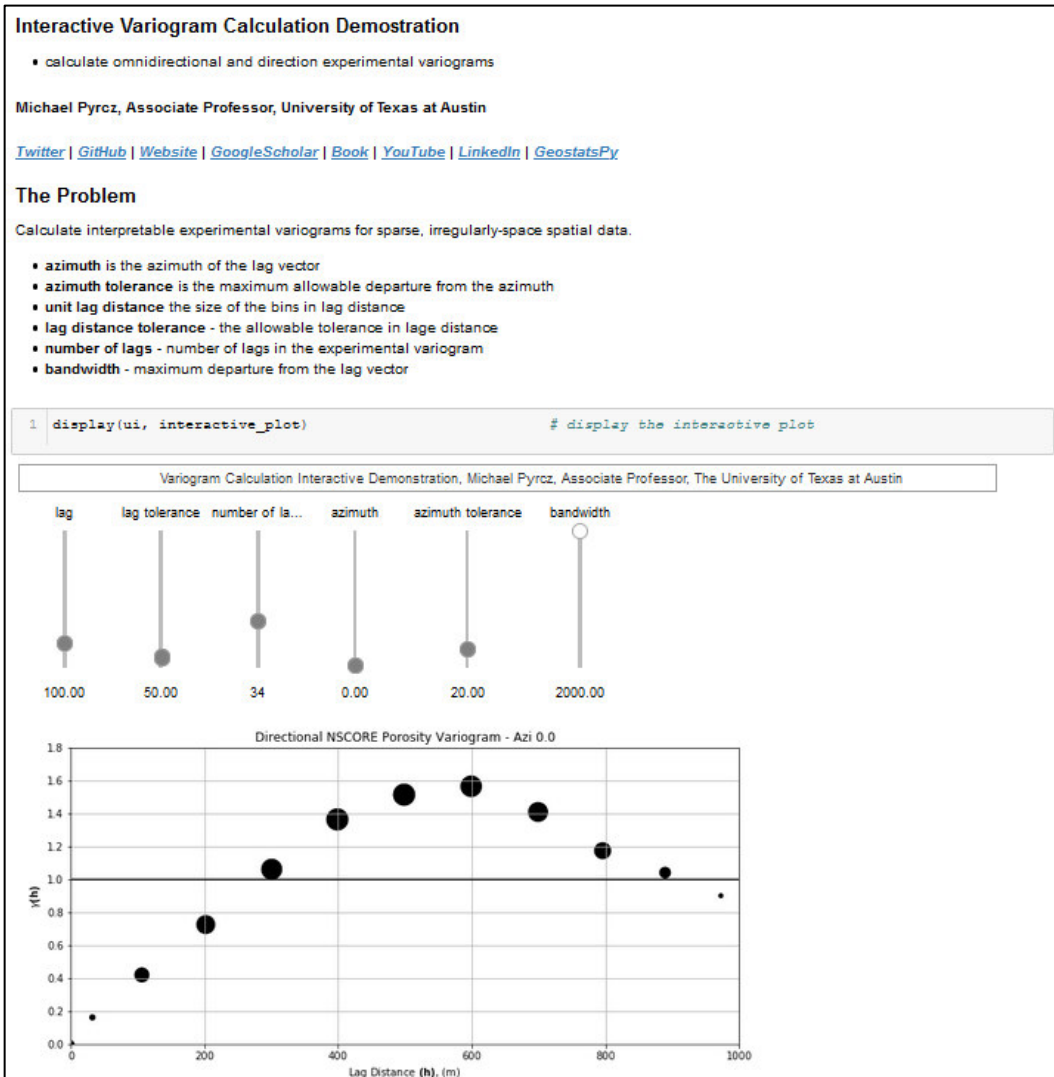
## Let's calculate variograms:

- normal score transformed (standard Gaussian) porosity

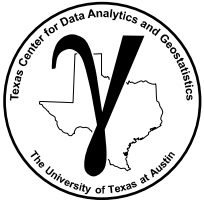
## Some Hints:

- start with lag of the minimum data spacing
- start with lag tolerance as  $\frac{1}{2}$  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).

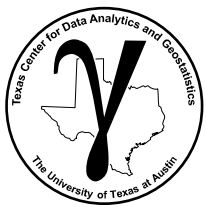


# Open Source Spatial Data Analytics in Python with GeostatsPy

## Introduction to GeostatsPy

Lecture outline . . .

- Workflow with GeostatsPy



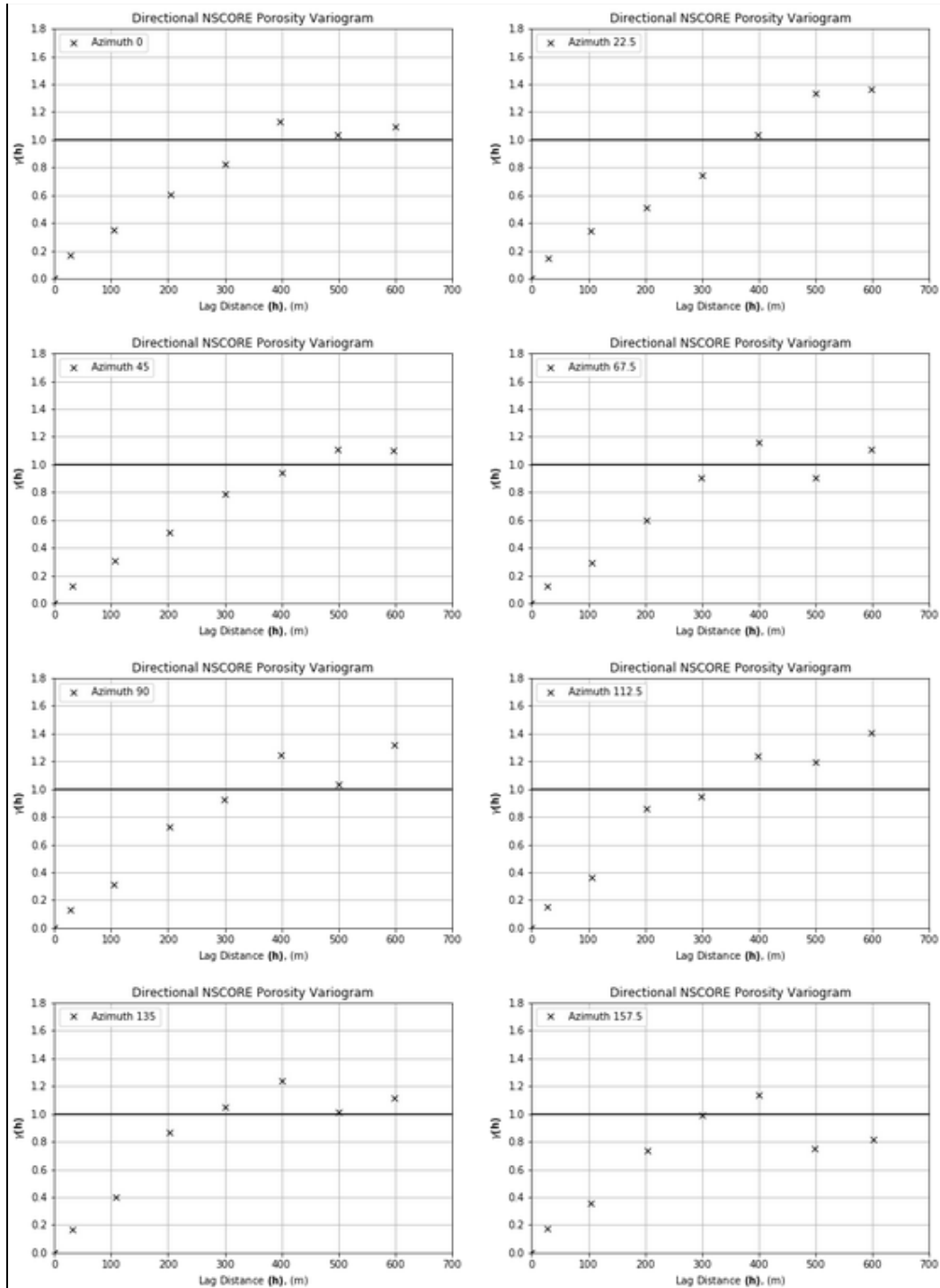
# Variogram Calculation Workflow with GeostatsPy

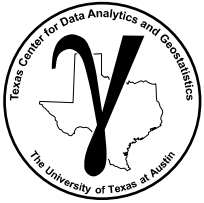
Let's walkthrough a more thorough variogram-based spatial analysis workflow:

- 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](#) | [Google Scholar](#) | [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 calculation with irregularly sampled data. This should help you get started with variogram calculation 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 are 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

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

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,  $C_s(\mathbf{h})$  and univariate variance,

$$C_s(\mathbf{h}) = \sigma_s^2 - \gamma(\mathbf{h})$$

the covariance function as:

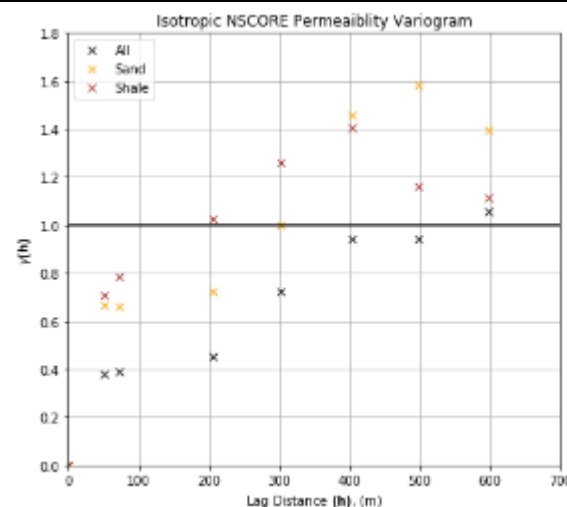
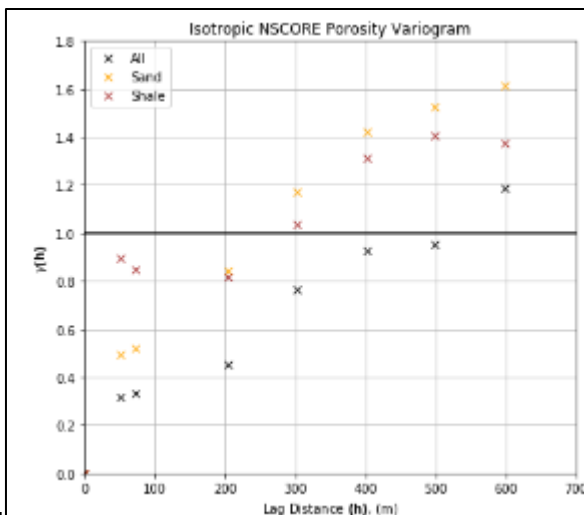
$$\rho_s(\mathbf{h}) = \frac{C_s(\mathbf{h})}{\sigma_s^2}$$

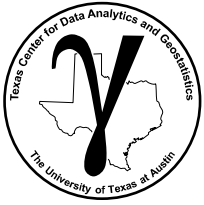
ion of the  $\mathbf{h} - \mathbf{h}$  scatter plot correlation vs. lag offset  $\mathbf{h}$ .

$$-1.0 \leq \rho_s(\mathbf{h}) \leq 1.0$$

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