Geostatistics and Machine Learning
Spatial Continuity and Prediction

Lecture outline . . .

- **Stationarity**
- **Spatial Continuity**
- **Variogram Calculation**
- **Spatial Estimation**

Instructor: Michael Pyrcz, the University of Texas at Austin

Other Resources

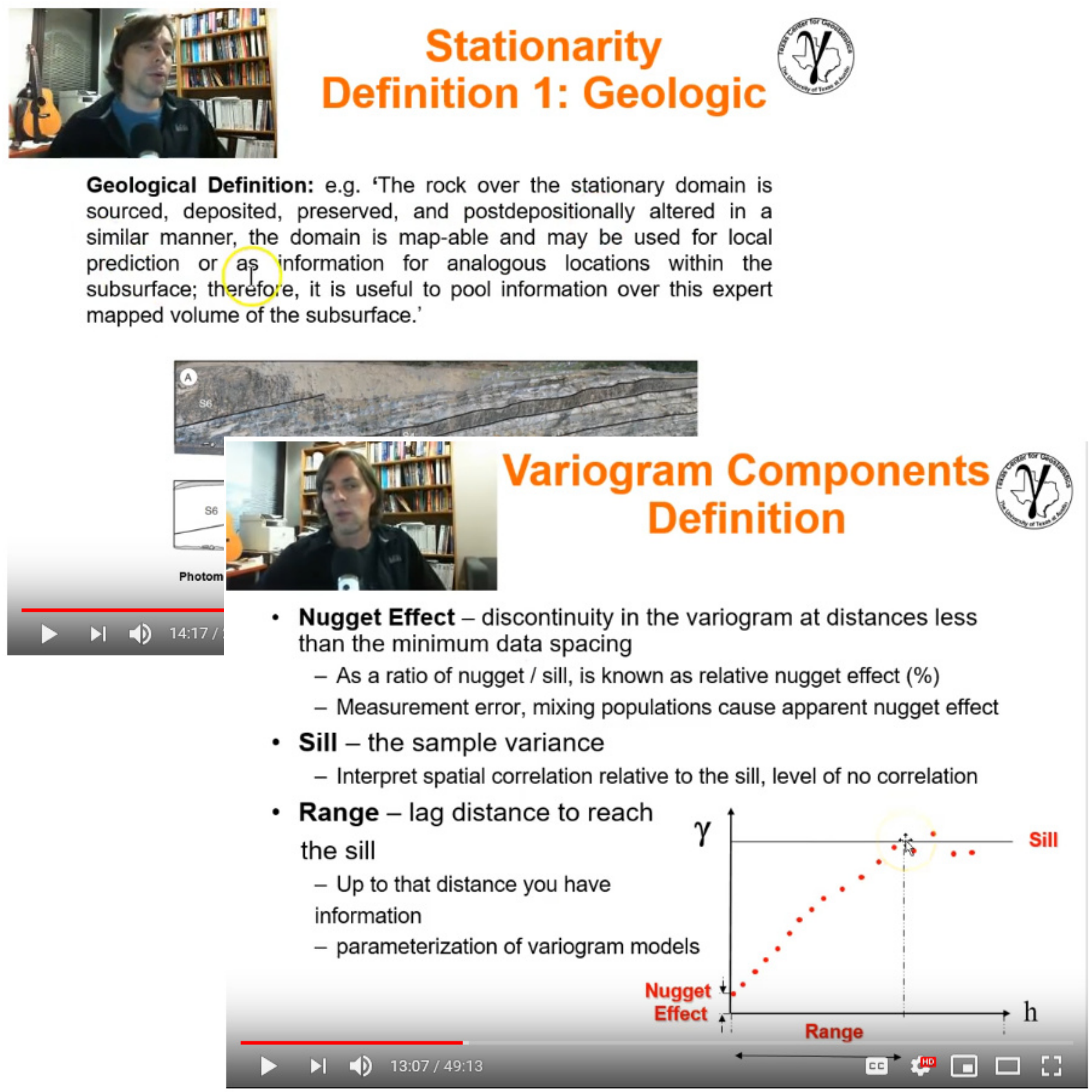
YouTube Lectures on:

Stationarity:
https://youtu.be/QwxQ9xuUHIU

Variogram Calculation:
https://youtu.be/mzPLicovE7Q

Variogram Interpretation:
https://youtu.be/Li-Xzlu7hvs

Variogram Modeling:
https://youtu.be/-Bi63Y3u6TU



Geostatistics and Machine Learning
Spatial Continuity and Prediction

Lecture outline . . .

- Stationarity

Introduction

Data Analytics

Inferential Methods

Predictive Methods

Advanced Methods

Conclusions

Instructor: Michael Pyrcz, the University of Texas at Austin

Stationarity Substituting Time for Space

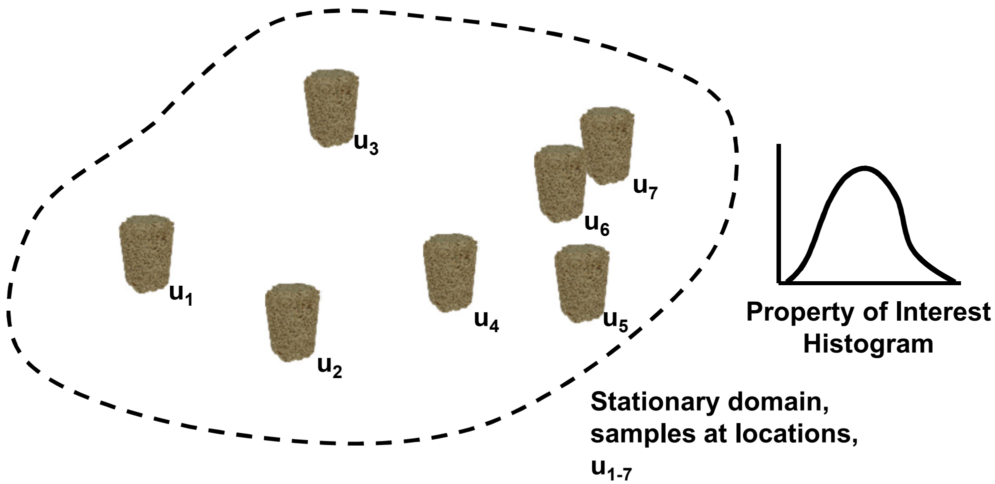
Any statistic requires replicates, repeated sampling (e.g. air or water samples from a monitoring station). In our geospatial problems repeated samples are not available at a location in the subsurface.

Extracted Sample at Location, u_1 Hole Remains at Location, u_1

Instead of time, **we must pool samples over space** to calculate our statistics. This decision to pool is the decision of stationarity. It is the decision that the subset of the subsurface is all the "same stuff".

Stationarity Substituting Time for Space

The decision of the stationary domain for sampling is an expert choice. Without it we are stuck in the “hole” and **cannot calculate any statistics** nor say anything about the behavior of the subsurface **between the sample data**.

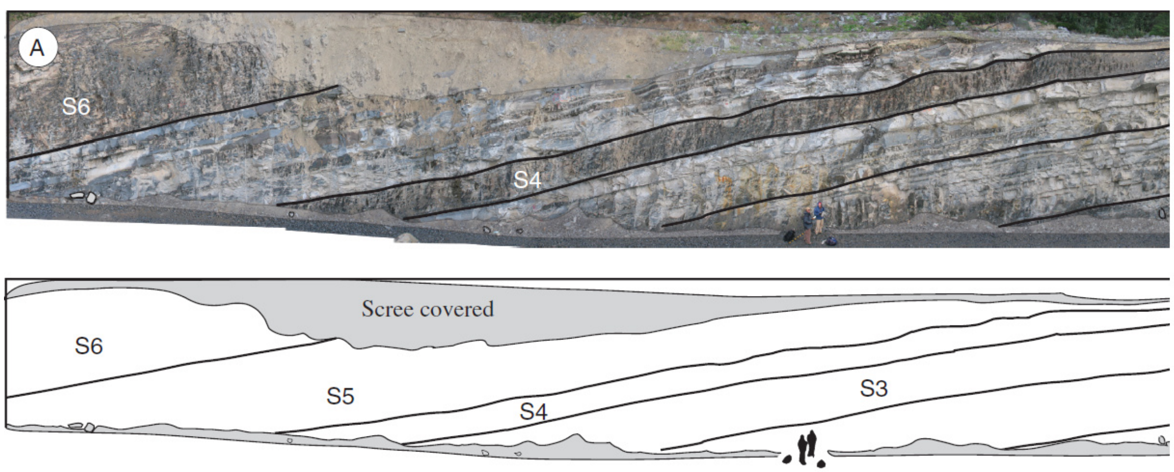


Import License: choice to pool specific samples to evaluate a statistic.

Export License: choice of where in the subsurface this statistic is applicable.

Stationarity
Definition 1: Geologic

Geological Definition: e.g. 'The rock over the stationary domain is sourced, deposited, preserved, and postdepositionally altered in a similar manner, the domain is map-able and may be used for local prediction or as information for analogous locations within the subsurface; therefore, it is useful to pool information over this expert mapped volume of the subsurface.'



Photomosaic, line drawing Punta Barrosa Formation sheet complex (Fildani et al. (2009).

Stationarity
Definition 2: Statistical

Statistical Definition: The metrics of interest are invariant under translation over the domain. For example, one point stationarity indicates the that histogram and associated statistics do not rely on location, **u**. Statistical stationarity for some common statistics:

Stationary Mean: $E\{Z(\mathbf{u})\} = m, \ \forall \, \mathbf{u}$

Stationary Distribution: $F(\mathbf{u}; z) = F(z), \ \forall \, \mathbf{u}$

Stationary Semivariogram: $\gamma_z(\mathbf{u}; \mathbf{h}) = \gamma_z(\mathbf{h}), \ \forall \, \mathbf{u}$

Stationarity: *What metric / statistic? Over what volume?*

May be extended to any statistic of interest including, facies proportions, bivariate distributions and multiple point statistics.

Geostatistics and Machine Learning
Spatial Continuity and Prediction

Lecture outline . . .

- **Spatial Continuity**

Introduction

Data Analytics

Inferential Methods

Predictive Methods

Advanced Methods

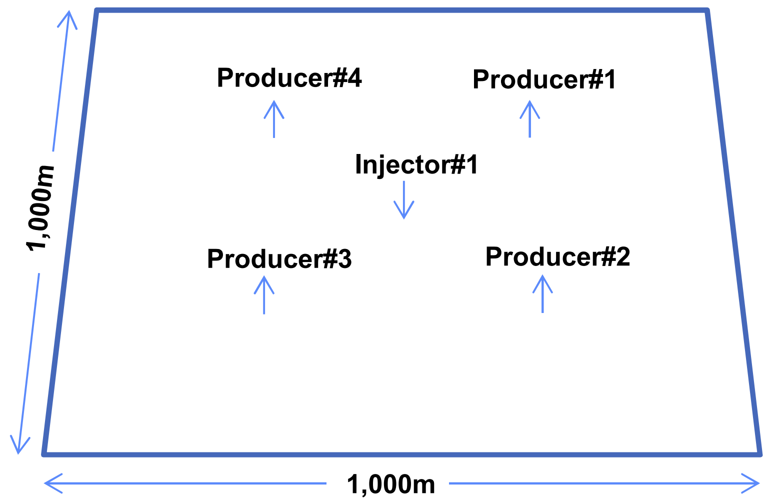
Conclusions

Instructor: Michael Pyrcz, the University of Texas at Austin

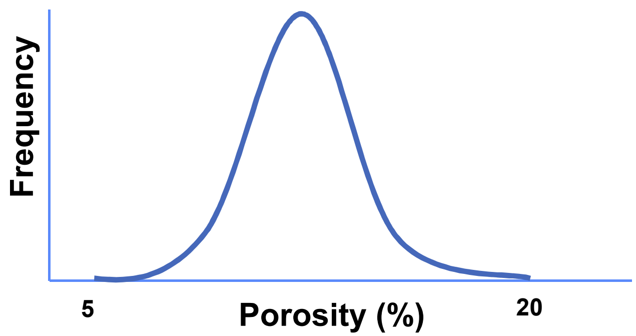
Motivation for Measuring Spatial Continuity

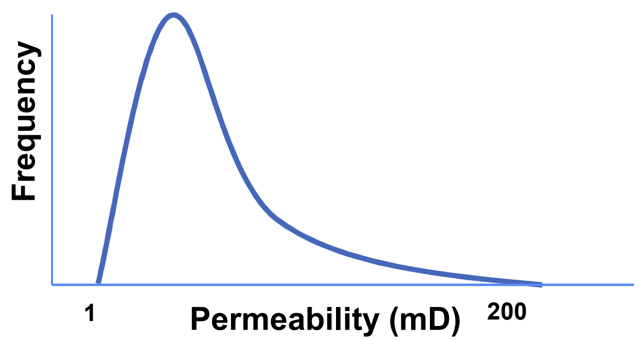
Simple Example

- Area of interest
- 1 Injector and 4 producers



- Porosity and permeability distributions (held constant for all cases)

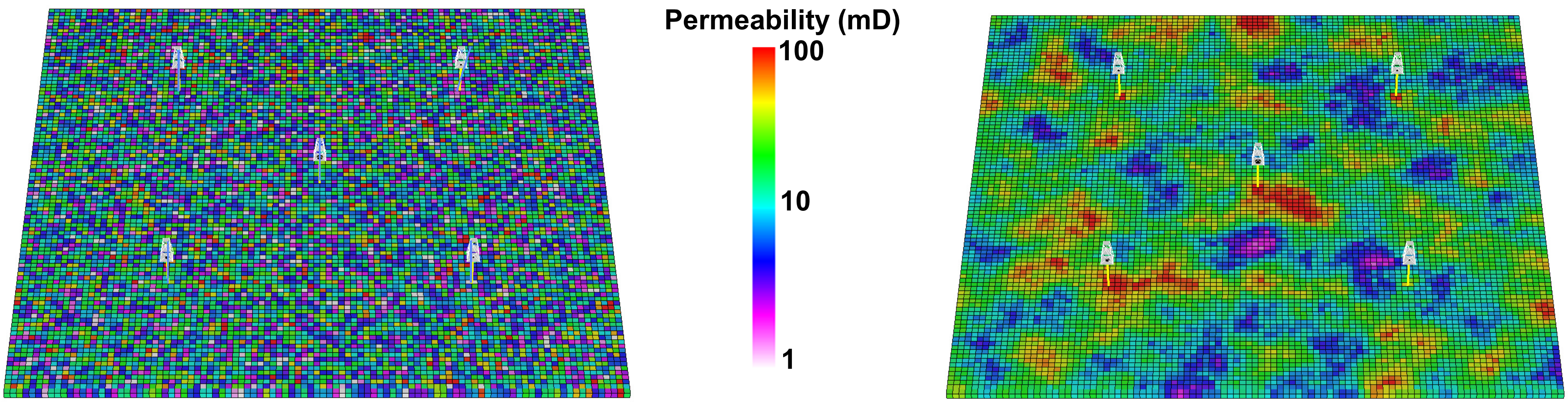




Motivation for Measuring Spatial Continuity

Does spatial continuity of reservoir properties matter?

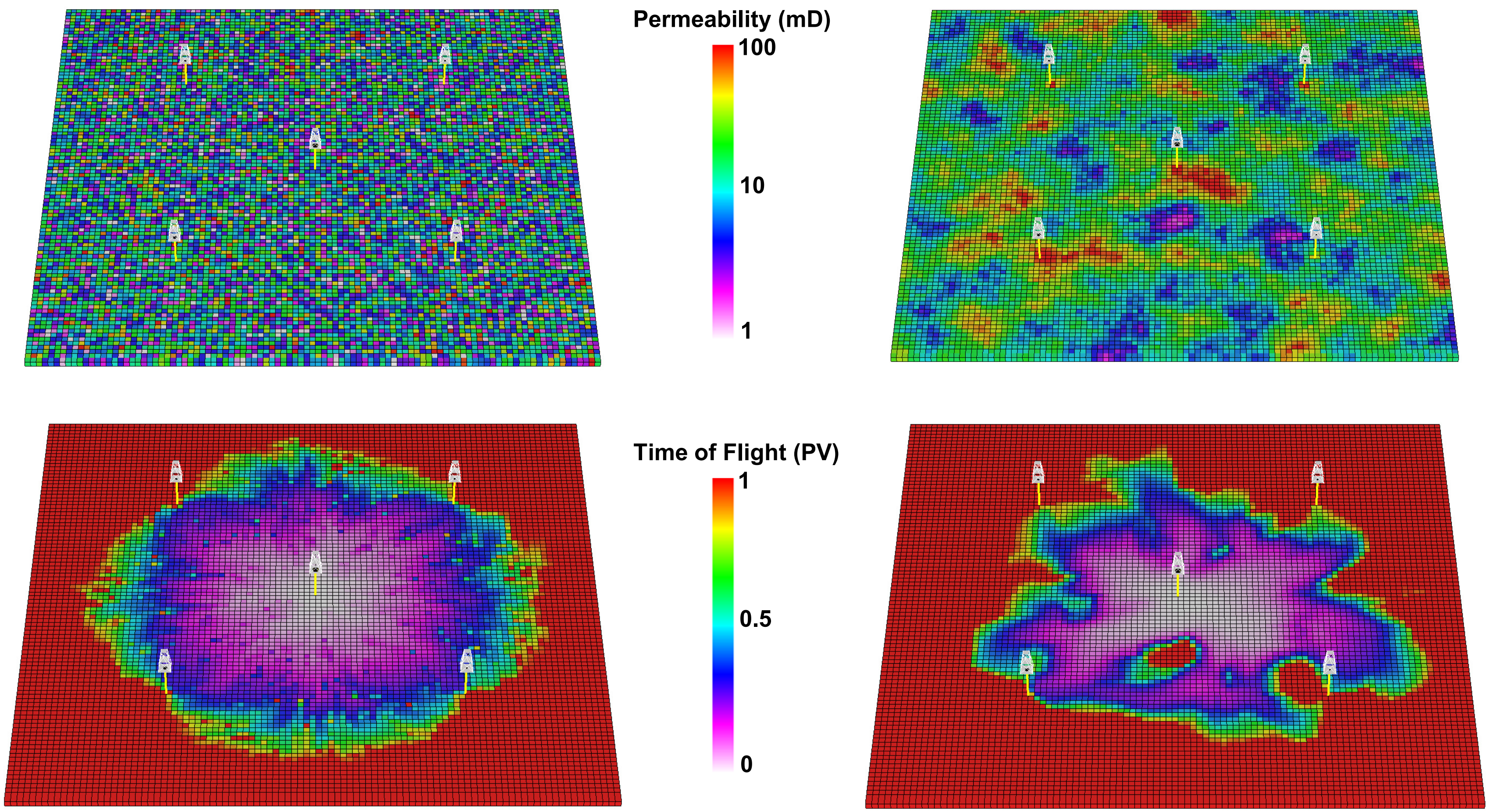
Consider these models of permeability



Recall – all models have the same porosity and permeability distributions

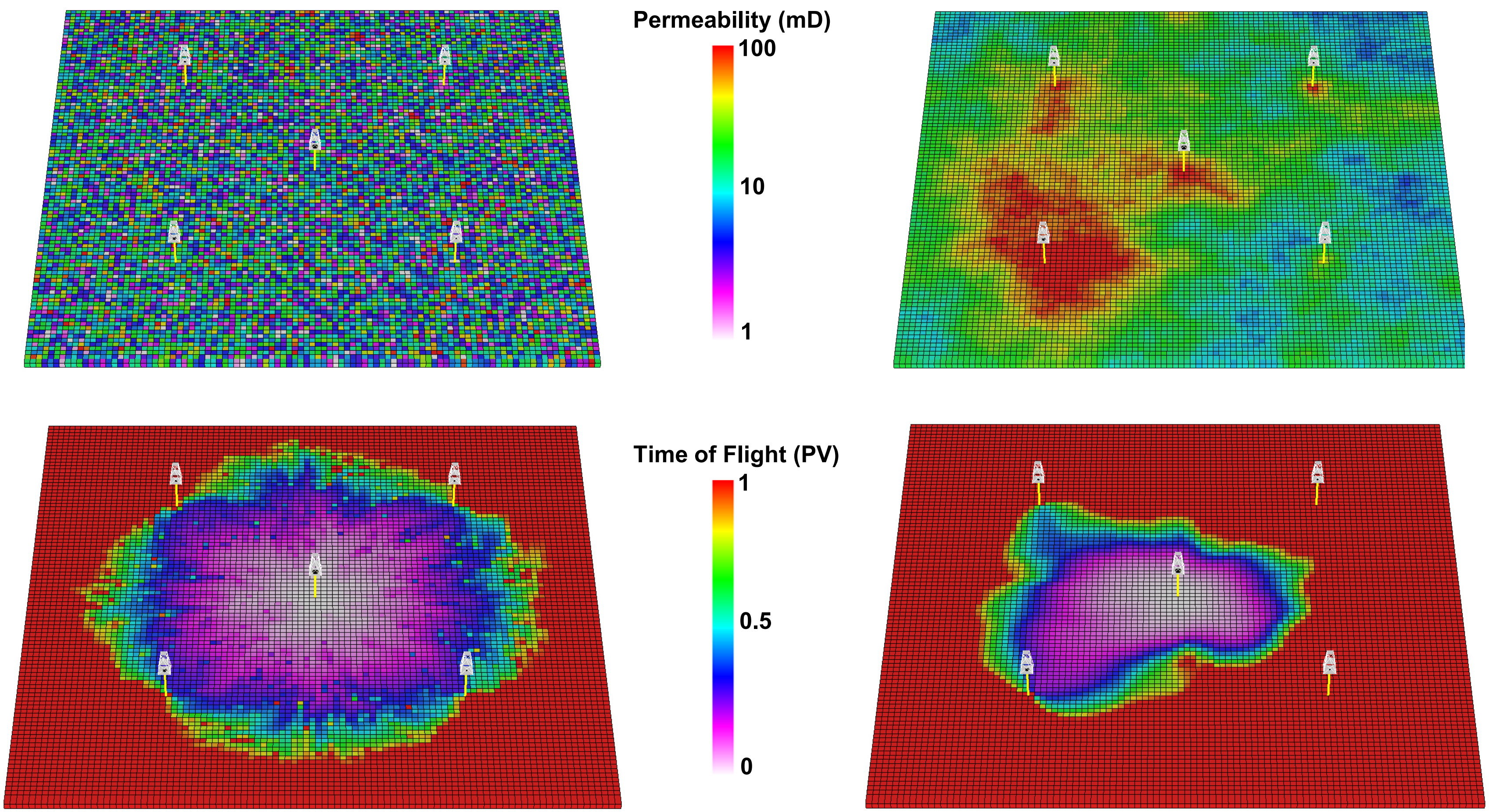
- Mean, variance, P10, P90 …
- Same static oil in place!

Motivation for Measuring Spatial Continuity



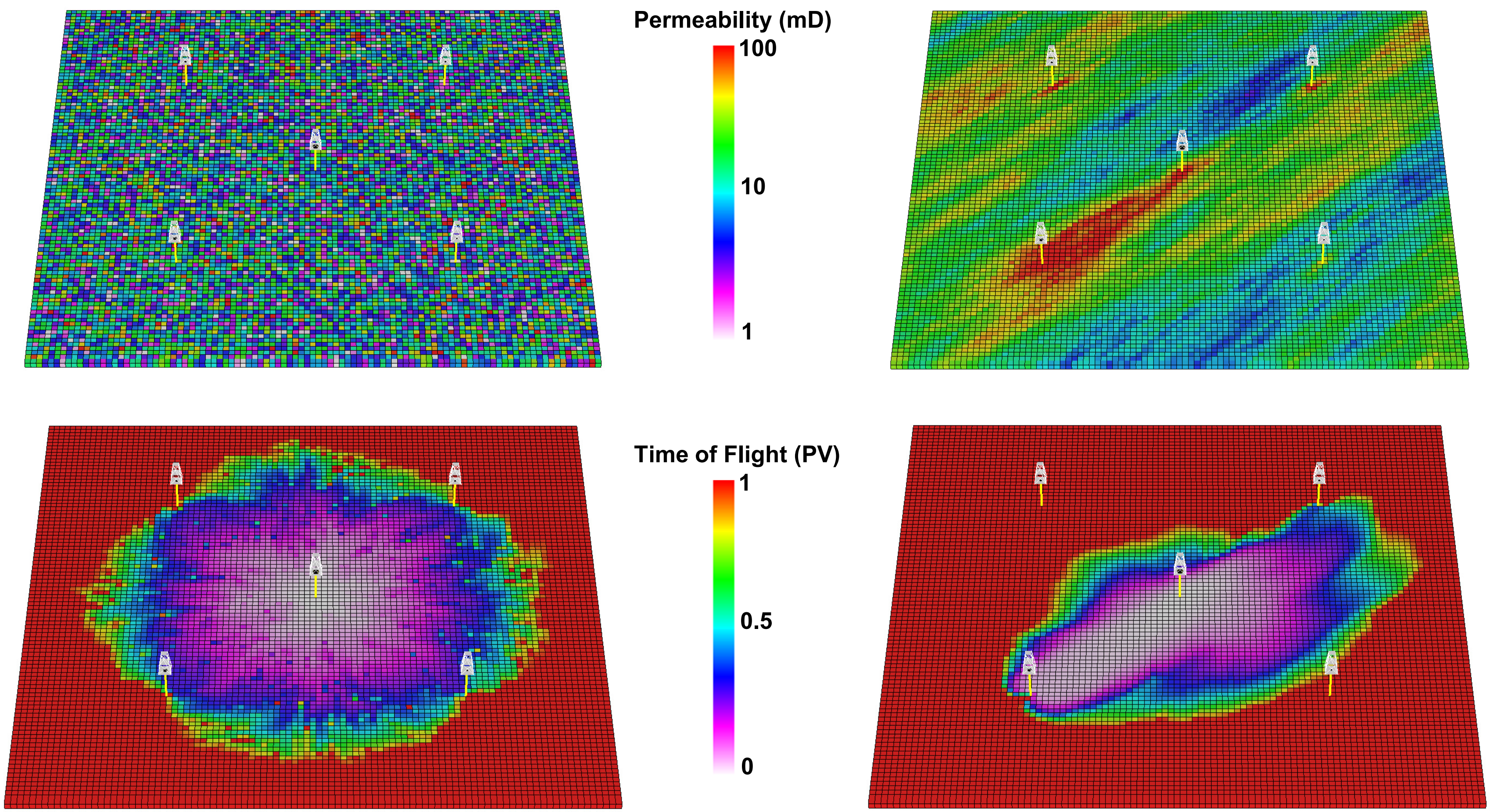
How does heterogeneity impact recovery factor?

Motivation for Measuring Spatial Continuity



How does heterogeneity impact recovery factor?

Motivation for Measuring Spatial Continuity



How does heterogeneity impact recovery factor?

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 Models

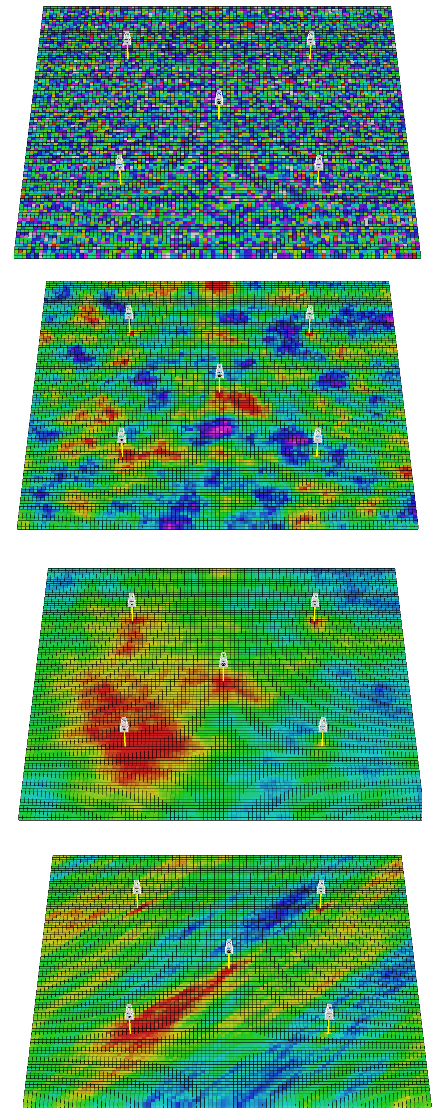
Spatial Continuity

"Very Short"

"Medium"

"Long"

"Anisotropic "



Spatial Continuity Definition

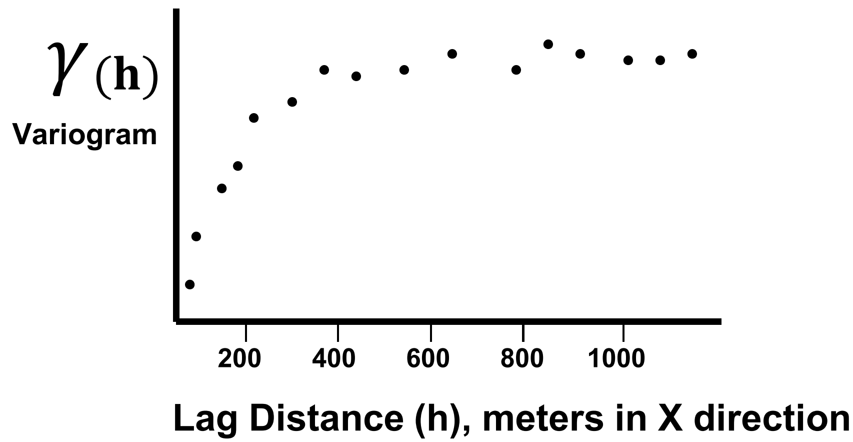
- **Spatial Continuity** – correlation between values over distance.
  - No spatial continuity – no correlation between values over distance, random values at each location in space regardless of separation distance.
  - Homogenous phenomenon have perfect spatial continuity, since all values as the same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

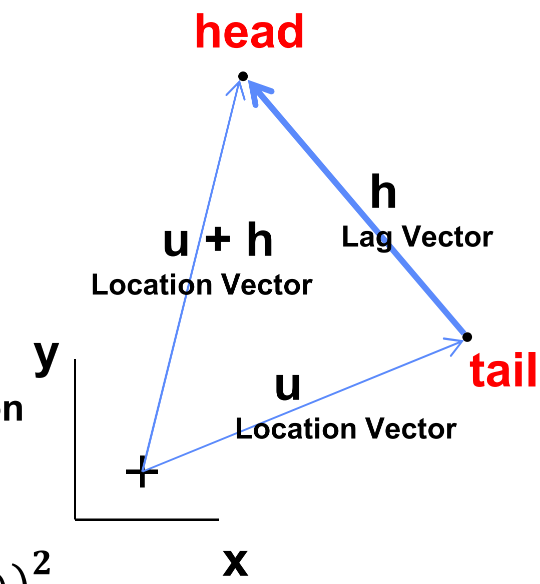
Measuring Spatial Continuity

We need a statistic to quantify spatial continuity!

The Semivariogram:

- Function of difference over distance.





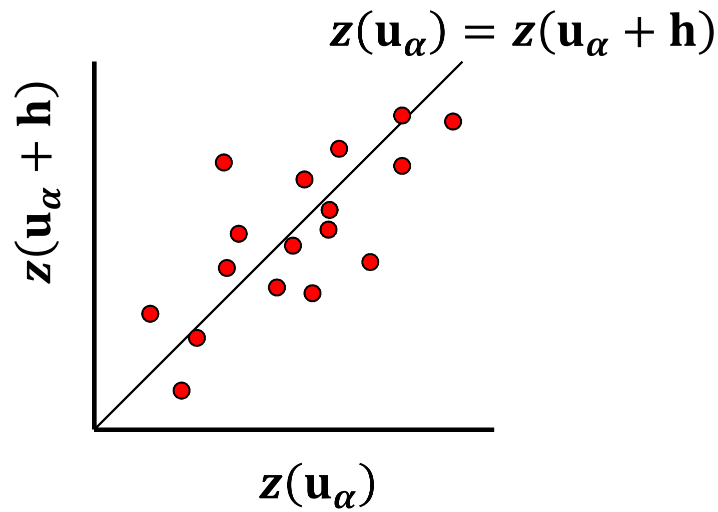
- The equation:

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

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

Modified from Deutsch, MinE310

“h” Scatterplot



- The variogram calculated for lag distance, $\mathbf{h}$, corresponds to the expected value of squared difference:

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

- Calculate for a suite of lag distances to obtain a continuous function.

Variogram Definition

- **Variogram** – a measure of dissimilarity vs. distance. Calculated as ½ the average squared difference of values separated by a lag vector.

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

  - The precise term is semivariogram (variogram if you remove the 1/2), but in practice the term variogram is used.
  - The ½ is used so that the covariance function and variogram may be related directly:

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

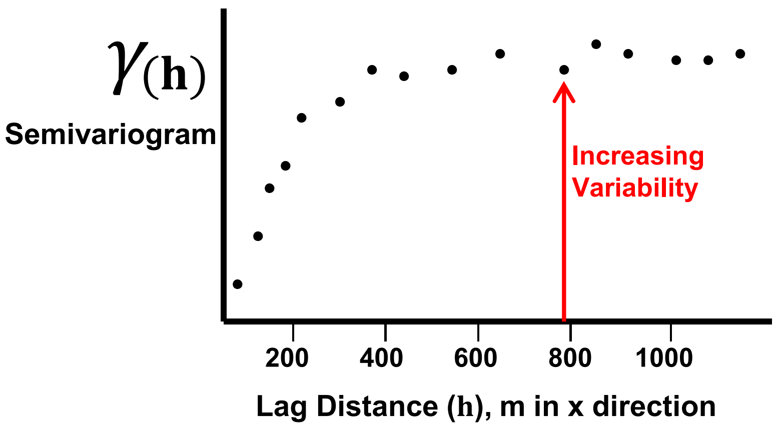
  - Note the correlogram is related to the covariance function as:

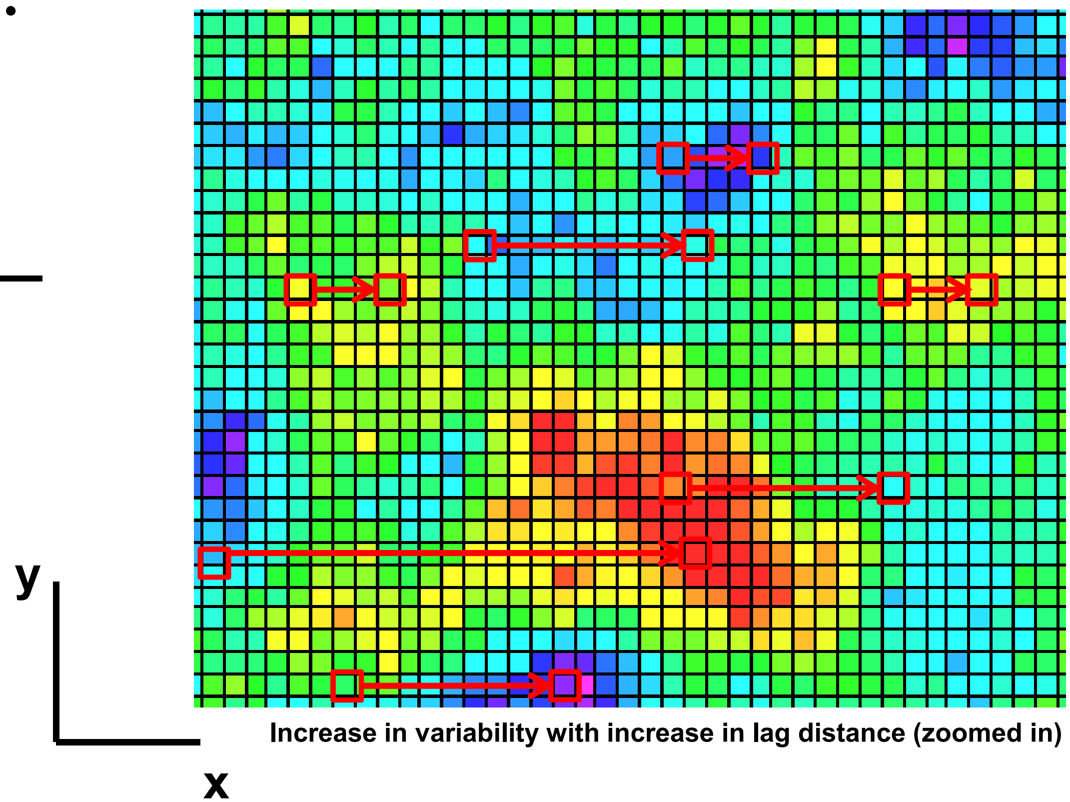
$$\rho_x(h) = \frac{C_x(h)}{\sigma_x^2}$$, h-scatter plot correlation vs. lag distance

Variogram Observations

Observation #1

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



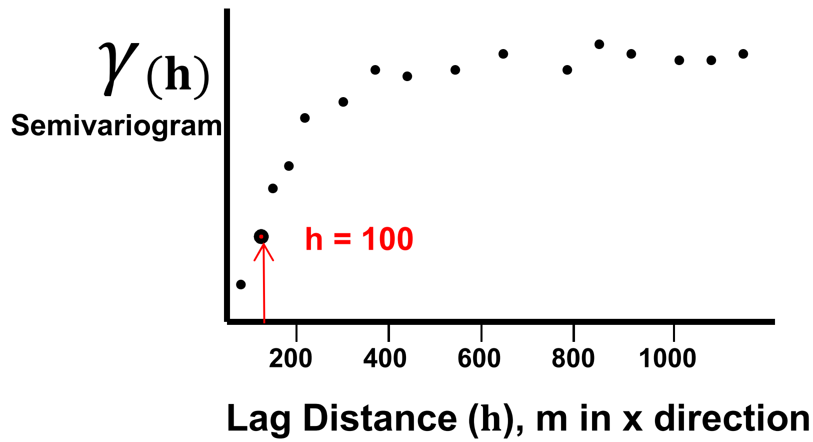


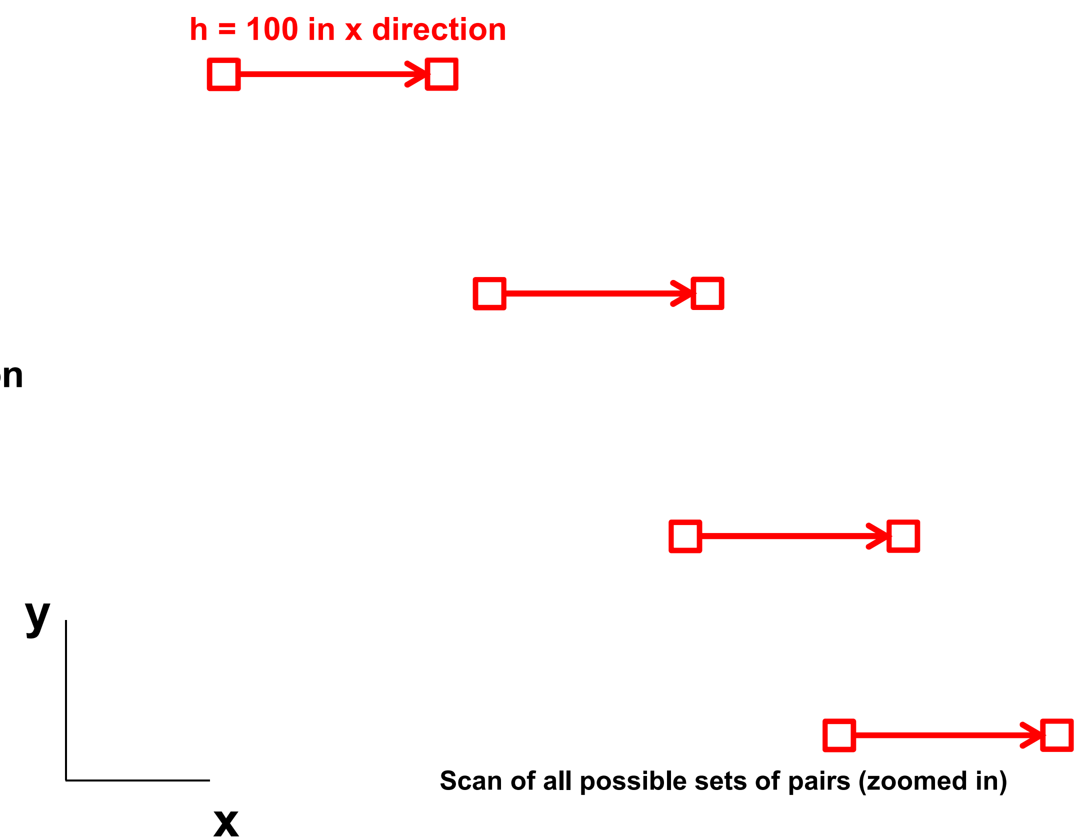
Increase in variability with increase in lag distance (zoomed in)

Variogram Observations

Observation #2

- Calculated with over all possible pairs separated by lag vector, $\mathbf{h}$.





- **The variogram:**

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

Given the number of pairs available N(h).

Variogram Observations

Observation #3

- Need to plot the sill to know the degree of correlation.

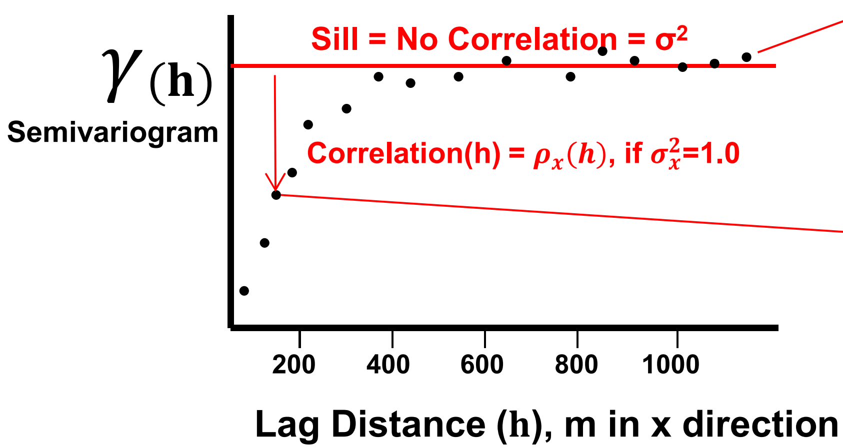
  Sill is the Variance, σ^2

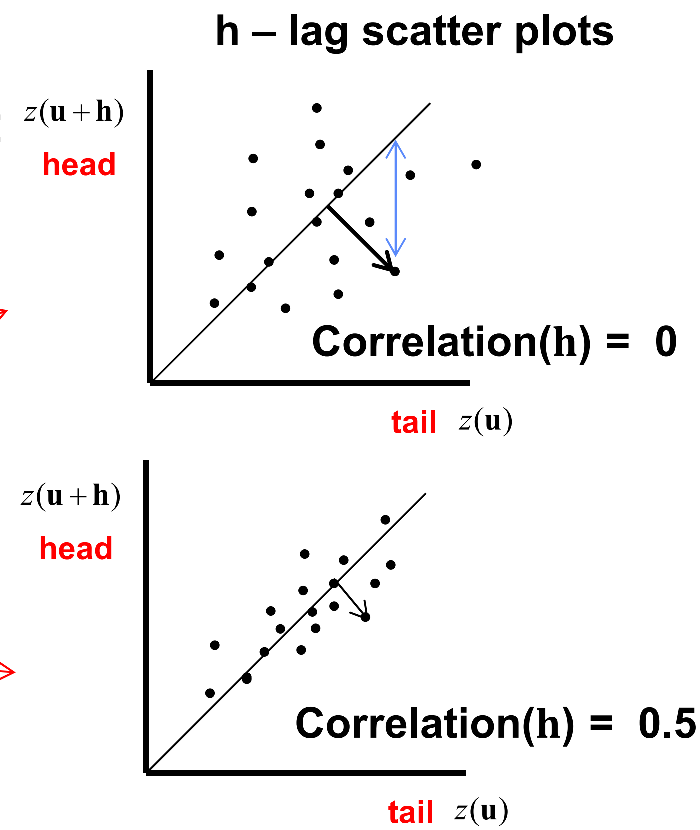
- Given stationarity of the variance and $\gamma_x(\mathbf{h})$:

Covariance Function: $C_x(h) = \sigma_x^2 - \gamma_x(\mathbf{h})$

- Given a standardized distribution $\sigma_x^2 = 1.0$:

Correlogram: $\rho_x(h) = \sigma_x^2 - \gamma_x(\mathbf{h})$



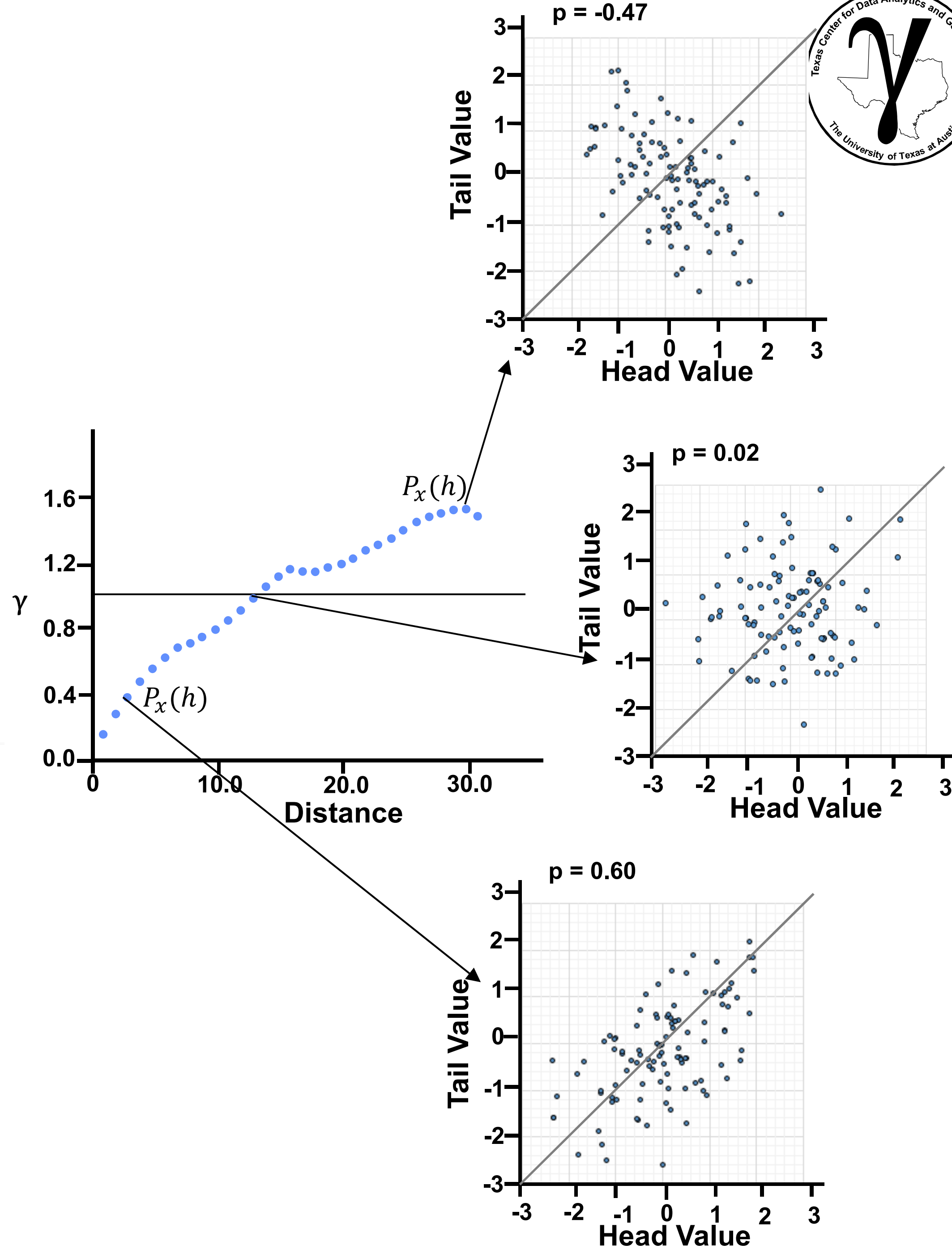


Variogram Interpretation

Observation #3

Need to plot the sill to know the degree of correlation.

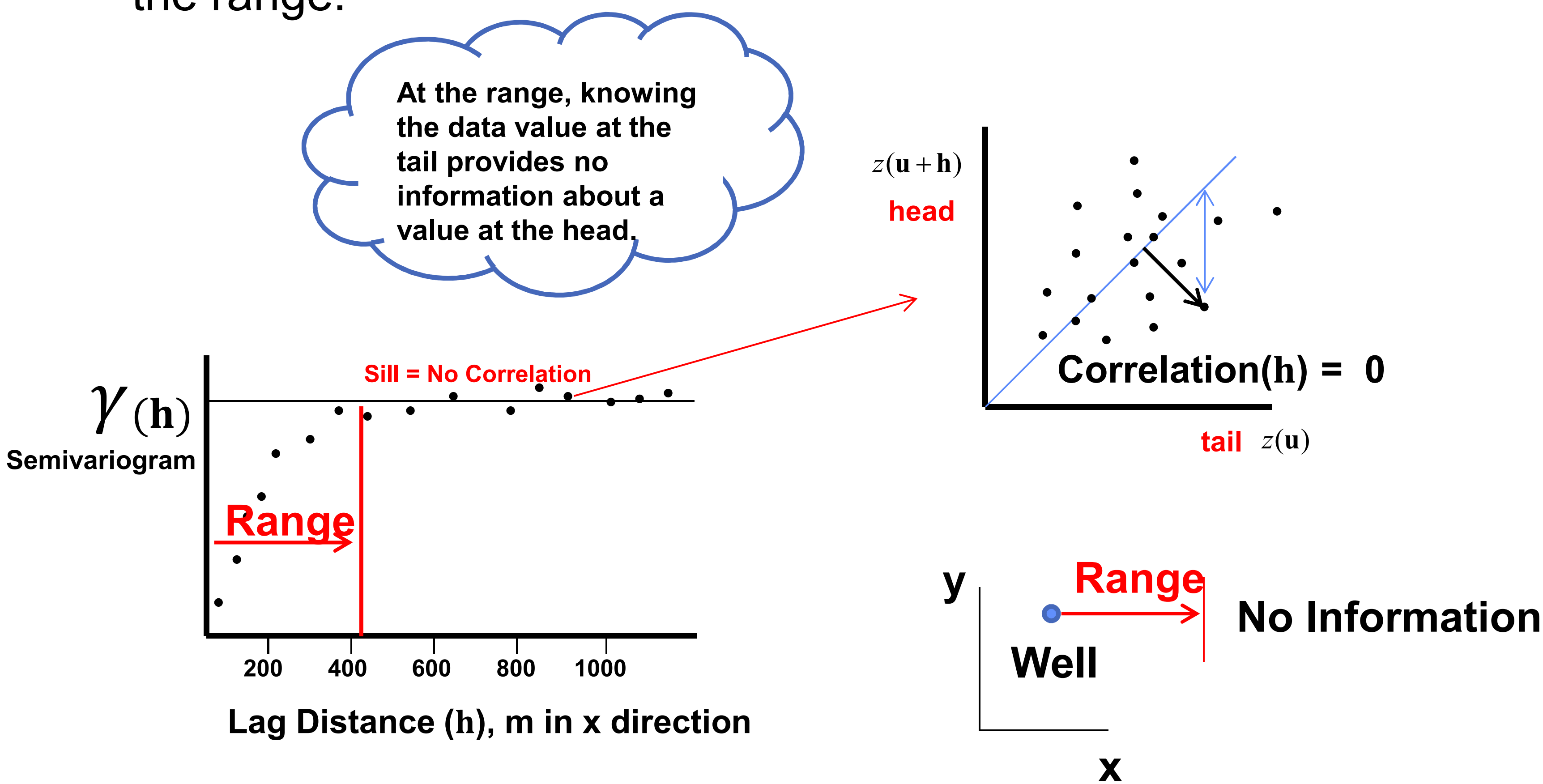
Another illustration of h-scatter plot correlation vs. lag distance.



Variogram Observations

Observation #4

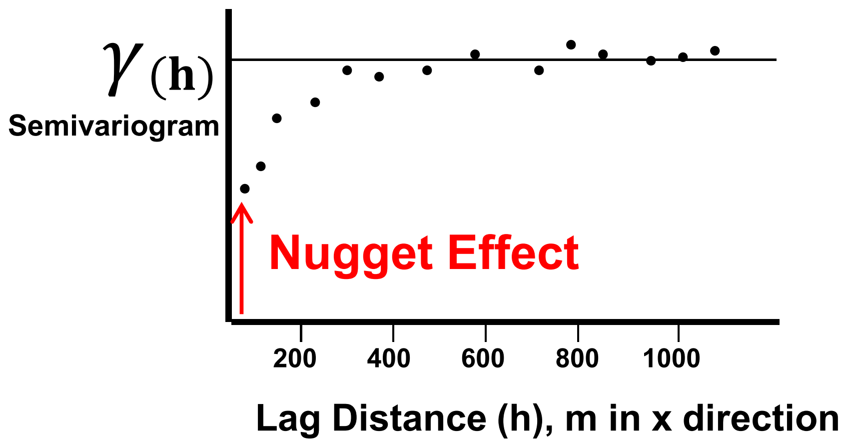
- The lag distance at which the variogram reaches the sill is know as the range.



Variogram Observations

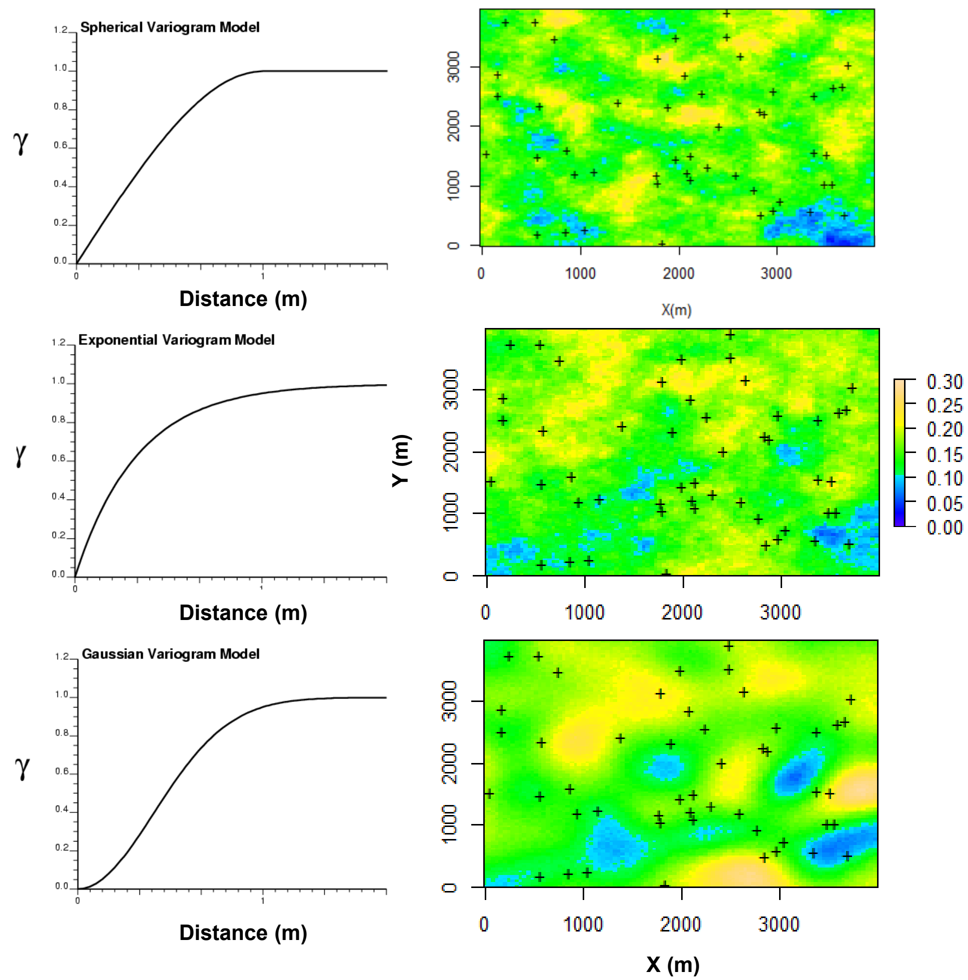
Observation #5

- Sometimes there is a discontinuity in the variogram at distances less than the minimum data spacing. This is known as nugget effect.
  - As a ratio of nugget / sill, is known as relative nugget effect (%)
  - Modeled as a no correlation structure that at lags, $h > \varepsilon$, an infinitesimal distance
  - Measurement error, mixing populations cause apparent nugget effect



Spatial Variability

- The three maps are remarkably similar: all three have the same 50 data, same histograms and same range of correlation, and yet their spatial variability/continuity is quite different
- The spatial variability/continuity depends on the detailed distribution of the petrophysical attribute (ϕ,K)
- The charts on the left are “variograms”
- Our map-making efforts should consider the spatial variability/continuity of the variable we are mapping:
  - Variability
  - Uncertainty



Porosity Realizations with 3 isotropic variograms.

Geostatistics and Machine Learning
Spatial Continuity and Prediction

Lecture outline . . .

- **Variogram Calculation**

Introduction

Data Analytics

Inferential Methods

Predictive Methods

Advanced Methods

Conclusions

Instructor: Michael Pyrcz, the University of Texas at Austin

Variogram Definition

- **Variogram** – a measure of dissimilarity vs. distance. Calculated as ½ the average squared difference of values separated by a lag vector.

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

  - The precise term is semivariogram (variogram if you remove the 1/2), but in practice the term variogram is used.
  - The ½ is used so that the covariance function and variogram may be related directly:

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

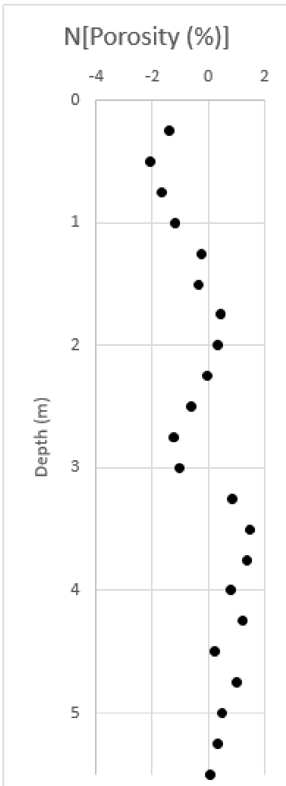
  - Note the correlogram is related to the covariance function as:

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

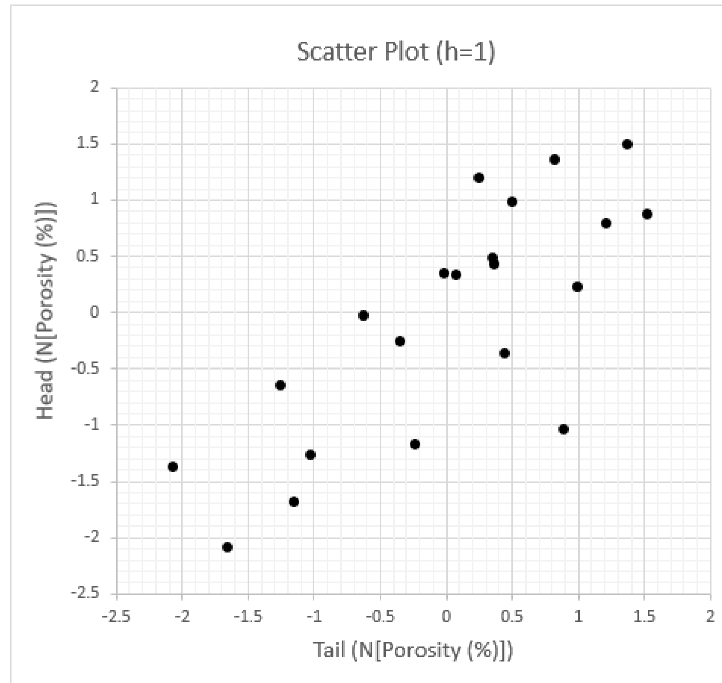
Variogram Calculation

- Consider data values separated by *lag* vectors (the h values)
- Here are two examples of a lag vector equal to the data spacing and then twice the data spacing:

Depth	N[Porosity]
0.25	-1.37
0.5	-2.08
0.75	-1.67
1	-1.16
1.25	-0.24
1.5	-0.36
1.75	0.44
2	0.36
2.25	-0.02
2.5	-0.63
2.75	-1.26
3	-1.03
3.25	0.88
3.5	1.51
3.75	1.37
4	0.81
4.25	1.21
4.5	0.24
4.75	0.99
5	0.49
5.25	0.34
5.5	0.07



Squared Difference	-1.37	
0.50	-2.08	-1.37
0.17	-1.67	-2.08
0.26	-1.16	-1.67
0.85	-0.24	-1.16
0.01	-0.36	-0.24
0.64	0.44	-0.36
0.01	0.36	0.44
0.14	-0.02	0.36
0.37	-0.63	-0.02
0.40	-1.26	-0.63
0.05	-1.03	-1.26
3.65	0.88	-1.03
0.40	1.51	0.88
0.02	1.37	1.51
0.31	0.81	1.37
0.16	1.21	0.81
0.94	0.24	1.21
0.56	0.99	0.24
0.25	0.49	0.99
0.02	0.34	0.49
0.07	0.07	0.34
Average / 2		0.07
0.23		

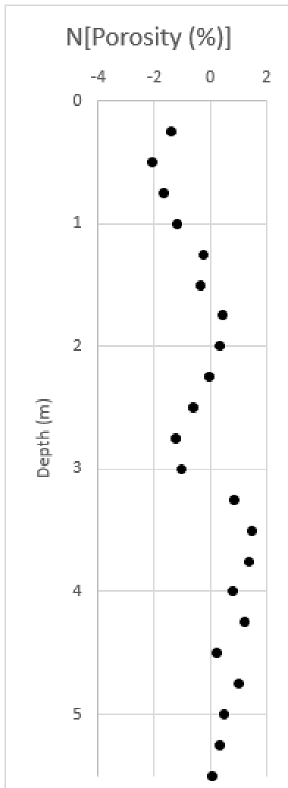


Correlation Coefficient (h=1)	0.77
Variogram (h=1)	0.23

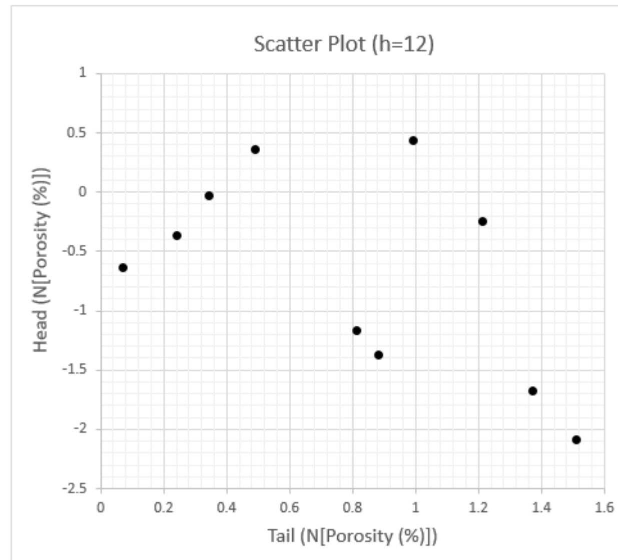
Variogram Calculation

- Consider data values separated by *lag* vectors (the h values)
- Here are two examples of a lag vector equal to the data spacing and then twice the data spacing:

Depth	N[Porosity]
0.25	-1.37
0.5	-2.08
0.75	-1.67
1	-1.16
1.25	-0.24
1.5	-0.36
1.75	0.44
2	0.36
2.25	-0.02
2.5	-0.63
2.75	-1.26
3	-1.03
3.25	0.88
3.5	1.51
3.75	1.37
4	0.81
4.25	1.21
4.5	0.24
4.75	0.99
5	0.49
5.25	0.34
5.5	0.07



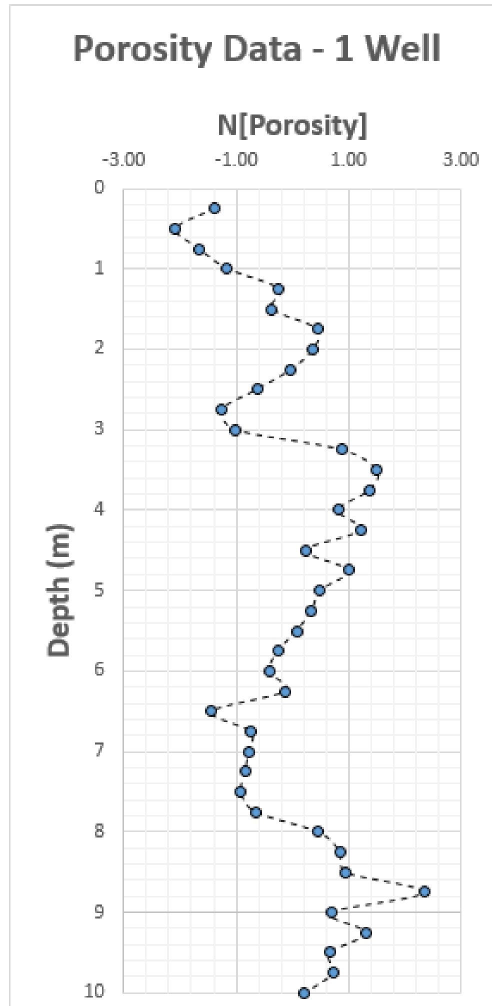
Squared Difference		
	-1.37	
	-2.08	
	-1.67	
	-1.16	
	-0.24	
	-0.36	
	0.44	
	0.36	
	-0.02	
	-0.63	
	-1.26	
	-1.03	
5.06	0.88	-1.37
12.89	1.51	-2.08
9.24	1.37	-1.67
3.88	0.81	-1.16
2.10	1.21	-0.24
0.36	0.24	-0.36
0.30	0.99	0.44
0.02	0.49	0.36
0.13	0.34	-0.02
0.49	0.07	-0.63
Average / 2		-1.26
1.72		-1.03
		0.88
		1.51
		1.37
		0.81
		1.21
		0.24
		0.99
		0.49
		0.34
		0.07



Correlation Coefficient (h=12)	-0.54
Variogram (h=12)	1.72

Variogram Calculation

- Pick a lag distance and calculate the variogram for that one lag distance.



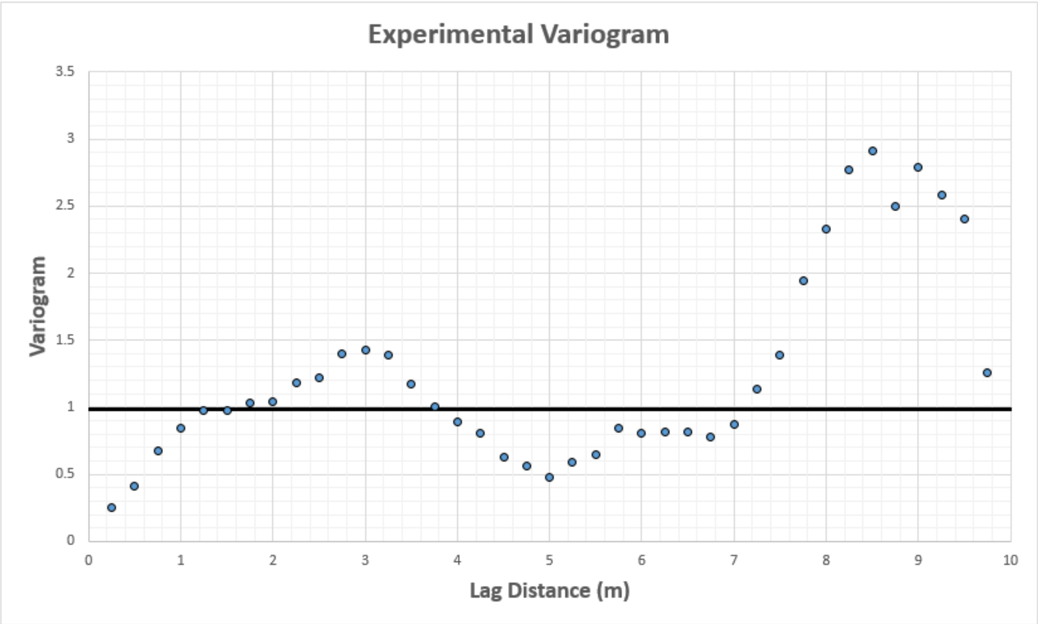
Depth	N[Porosity]
0.25	-1.37
0.5	-2.08
0.75	-1.67
1	-1.16
1.25	-0.24
1.5	-0.36
1.75	0.44
2	0.36
2.25	-0.02
2.5	-0.63
2.75	-1.26
3	-1.03
3.25	0.88
3.5	1.51
3.75	1.37
4	0.81
4.25	1.21
4.5	0.24
4.75	0.99
5	0.49
5.25	0.34
5.5	0.07
5.75	-0.26
6	-0.41
6.25	-0.14
6.5	-1.44
6.75	-0.75
7	-0.78
7.25	-0.85
7.5	-0.92
7.75	-0.66
8	0.47
8.25	0.85
8.5	0.95
8.75	2.35
9	0.69
9.25	1.31
9.5	0.66
9.75	0.72
10	0.21

The file is Variogram_Simple_Calculation.xlsx

Variogram Calculation Example

- Pick a lag distance and calculate the variogram for that one lag distance.

- Here's all of them:



Depth	N[Porosity]
0.25	-1.37
0.5	-2.08
0.75	-1.67
1	-1.16
1.25	-0.24
1.5	-0.36
1.75	0.44
2	0.36
2.25	-0.02
2.5	-0.63
2.75	-1.26
3	-1.03
3.25	0.88
3.5	1.51
3.75	1.37
4	0.81
4.25	1.21
4.5	0.24
4.75	0.99
5	0.49
5.25	0.34
5.5	0.07
5.75	-0.26
6	-0.41
6.25	-0.14
6.5	-1.44
6.75	-0.75
7	-0.78
7.25	-0.85
7.5	-0.92
7.75	-0.66
8	0.47
8.25	0.85
8.5	0.95
8.75	2.35
9	0.69
9.25	1.31
9.5	0.66
9.75	0.72
10	0.21

The Variogram and Covariance Function

- **The variogram, covariance function and correlation coefficient are equivalent tools for characterizing spatial two-point correlation** (assuming stationarity):

$$\gamma_x(\mathbf{h}) = \sigma_x^2 - C_x(h)$$
$$= \sigma_x^2(1 - \rho_x(h))$$

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

where:

$$C_x(\mathbf{h}) = E\{X(\mathbf{u}) \cdot X(\mathbf{u}+\mathbf{h})\} - [E\{X(\mathbf{u})\} \cdot E\{X(\mathbf{u}+\mathbf{h})\}], \forall\ \mathbf{u}, \mathbf{u}+\mathbf{h} \in A$$

$$C_x(0) = \sigma_x^2$$

$$C_x(\mathbf{h}) = \frac{\sum_{\alpha=1}^{n} x(\mathbf{u}_\alpha) \cdot x(\mathbf{u}_\alpha + \mathbf{h})}{n} - (\overline{x})^2$$, if stationary mean

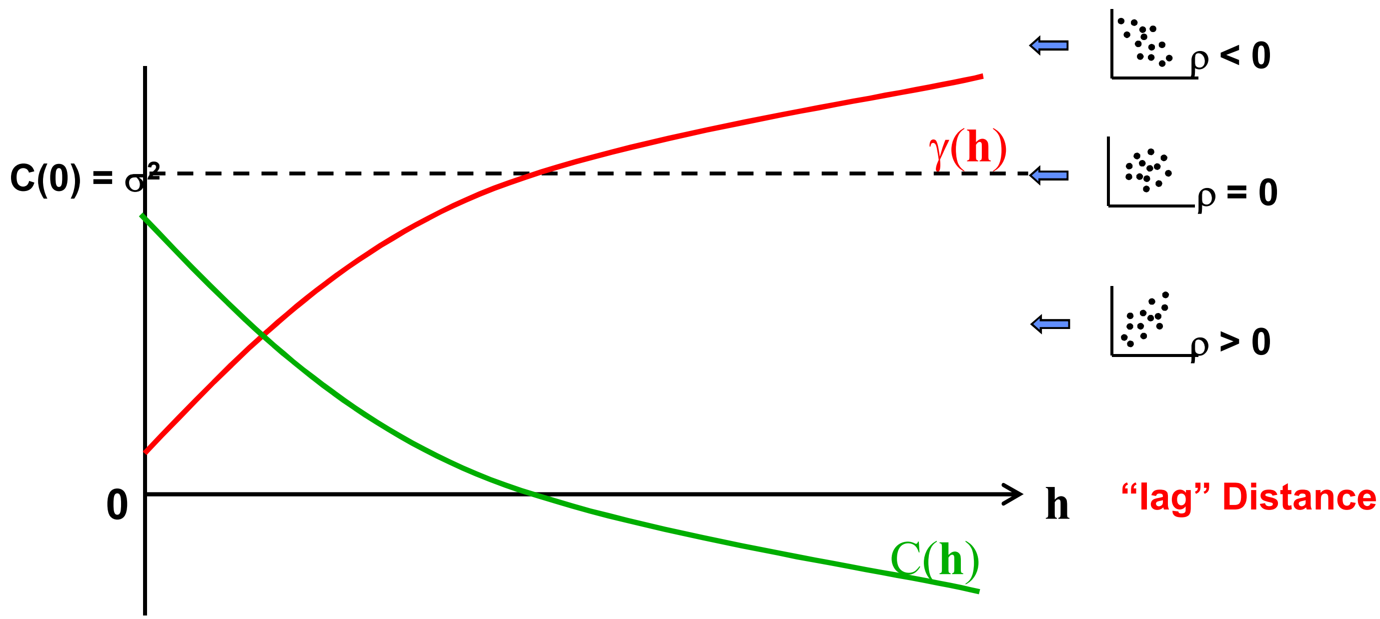
Stationarity entails that:

$$m(\mathbf{u}) = m(\mathbf{u}+\mathbf{h}) = m = E\{Z\},\ \forall \mathbf{u} \in A$$

$$Var(\mathbf{u}) = Var(\mathbf{u}+\mathbf{h}) = \sigma^2 = Var\{Z\},\ \forall \mathbf{u} \in A$$

The Variogram and Covariance Function

- **Must plot variance to interpret variogram:**
  - Positive correlation when semivariogram less than variance
  - No correlation when the semivariogram is equal to the variance
  - Negative correlation when the semivariogram points above variance



Covariance Function Definition

- **Covariance Function** – a measure of similarity vs. distance. Calculated as the average product of values separated by a lag vector centered by the square of the mean.

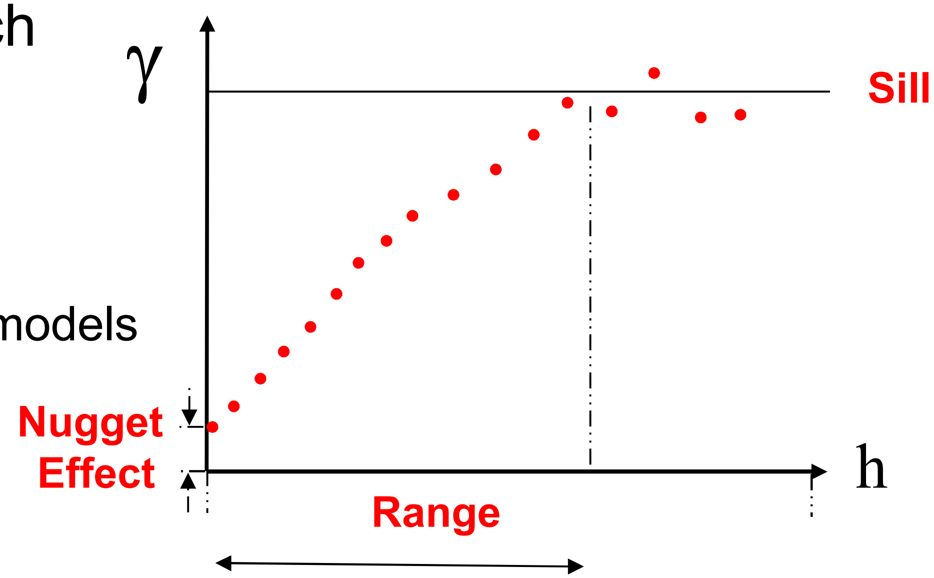
$$\boldsymbol{C_x}(\mathbf{h}) = \frac{\sum_{\alpha=1}^{n} x(\mathbf{u}_\alpha)\cdot x(\mathbf{u}_\alpha+\mathbf{h})}{\boldsymbol{n}} - (\overline{\boldsymbol{x}})^2, \text{ if stationary mean}$$

  – The covariance function is the variogram upside down. $\boldsymbol{\gamma_x}(\mathbf{h}) = \boldsymbol{\sigma_x^2} - \boldsymbol{C_x}(\boldsymbol{h})$

  – We model variograms, but inside the kriging and simulation methods they are converted to covariance values for numerical convenience.

Variogram Components Definition

- **Nugget Effect** – discontinuity in the variogram at distances less than the minimum data spacing
  - As a ratio of nugget / sill, is known as relative nugget effect (%)
  - Measurement error, mixing populations cause apparent nugget effect
- **Sill** – the sample variance
  - Interpret spatial correlation relative to the sill, level of no correlation
- **Range** – lag distance to reach the sill
  - Up to that distance you have information
  - parameterization of variogram models



Spoiler Alert

We need to practically calculate and model spatial continuity. From the available and often sparse subsurface data.

1. Calculate variogram with irregularly spaced data
   – Search templates with parameters

2. Valid spatial model
   – Fit with a couple different, nest (additive) spatial continuity models e.g. nugget, spherical, exponential and Gaussian

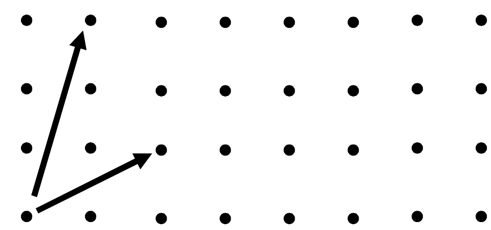
3. Full 3D spatial continuity model
   – Model primary directions, i.e. major horizontal, minor horizontal and vertical and combine together with assumption of geometric anisotropy

We will not cover variogram modeling.

Calculating Experimental Variograms

How do we get pairs separated by lag vector?

- Regular spaced data:
  - Specify as offsets of grid units
  - Fast calculation
  - Diagonal directions are awkward
- Irregular spaced data:
  - Nominal distance for each lag
  - Distance tolerance
  - Azimuth direction
  - Azimuth tolerance
  - Dip direction
  - Dip tolerance
  - Bandwidth (maximum deviation) in originally horizontal plane
  - Bandwidth in originally vertical plane



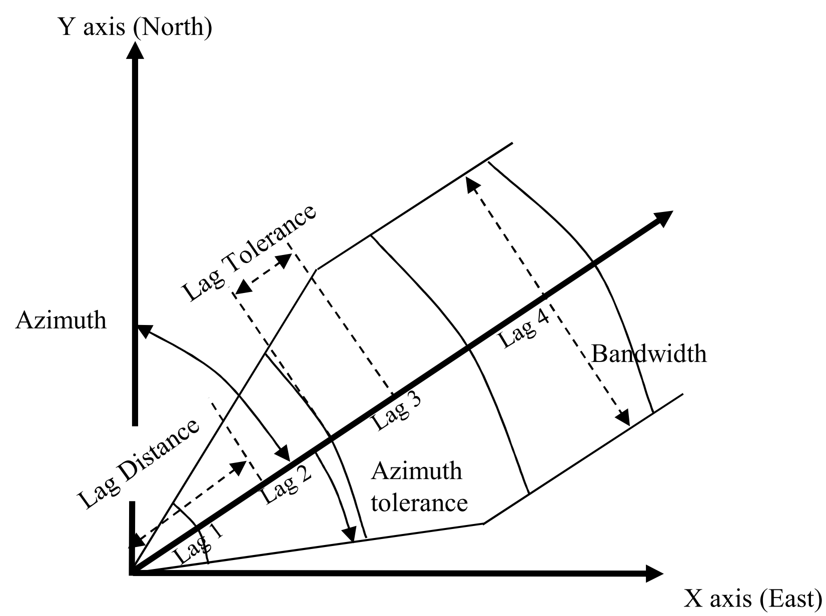
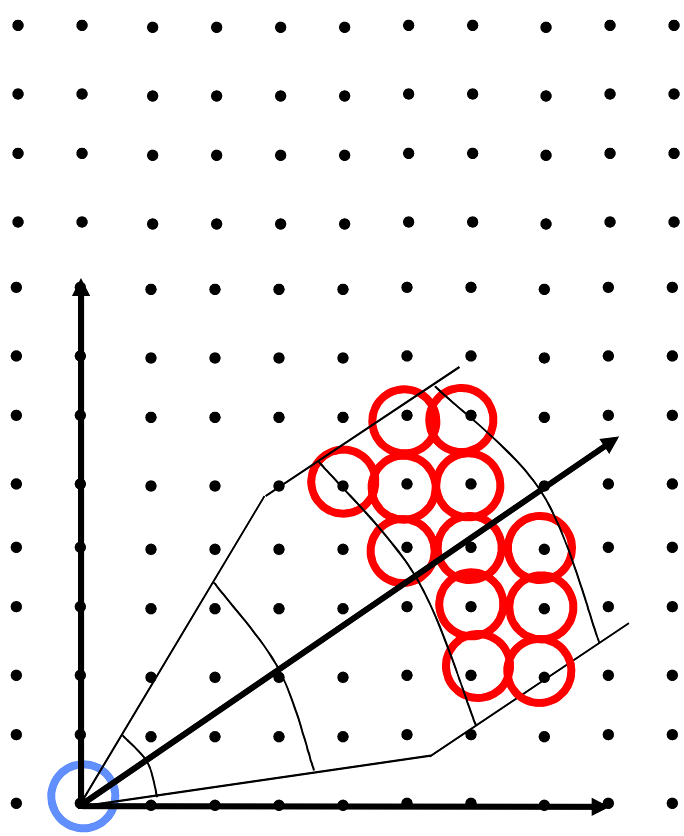


Image from Pyrcz and Deutsch, 2014

Calculating Experimental Variograms

Example: Starting With One Lag (i.e. #4)

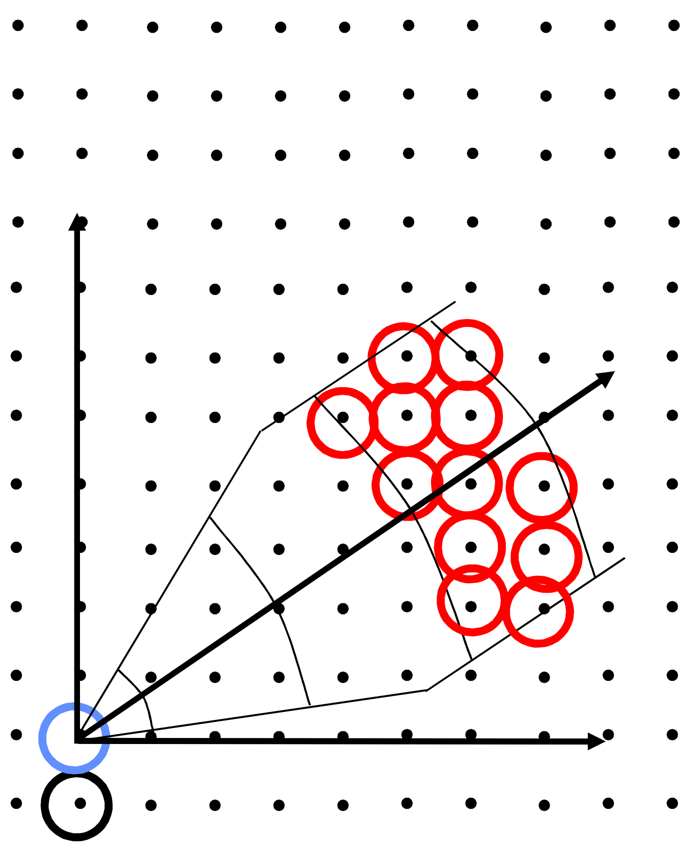


$$2\gamma(\boldsymbol{h}) = \frac{1}{N(\boldsymbol{h})} \sum_{N(\boldsymbol{h})} [z(\boldsymbol{u}) - z(\boldsymbol{u}+\boldsymbol{h})]^2$$

Start at a node, and compare value to all nodes which fall in the lag and angle tolerance.

...

Calculating Experimental Variograms



$$2\gamma(\boldsymbol{h}) = \frac{1}{N(\boldsymbol{h})} \sum_{N(\boldsymbol{h})} [z(\boldsymbol{u}) - z(\boldsymbol{u}+\boldsymbol{h})]^2$$

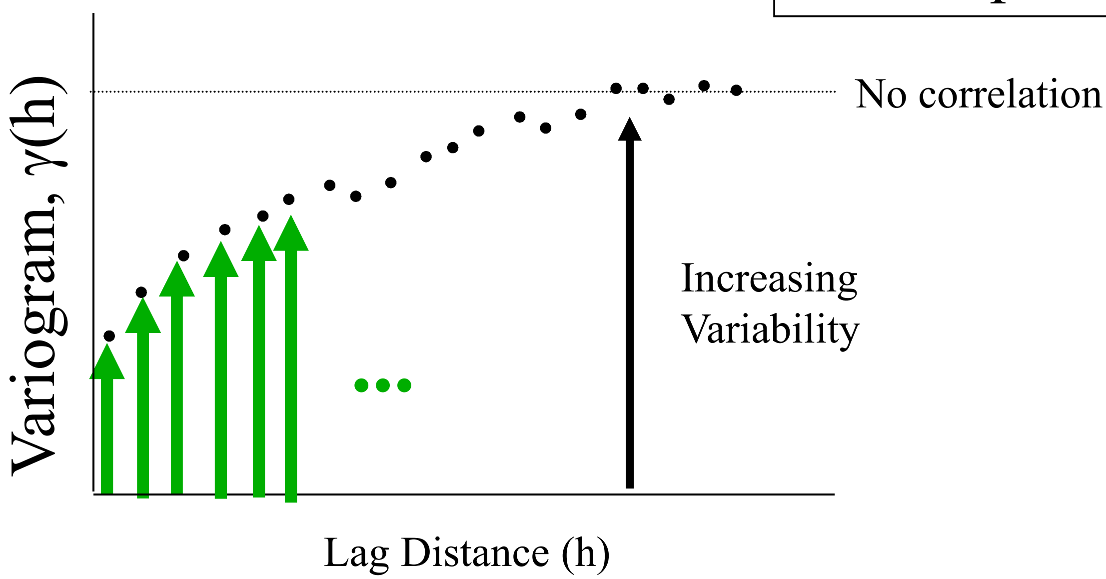
Move to next node.

...

Calculating Experimental Variograms

Now Repeat for All Nodes

And Repeat for All Lags



Some Options

- Data transformation:
  - Transform a continuous variable to a Gaussian or normal distribution, for use with / consistency with Gaussian simulation methods
  - Transform a categorical variable to a series of indicator variables for indicator methods and categorical to continuous for truncated Gaussian methods
- Coordinate transformation:
  - Variograms are calculated aligned with the stratigraphic framework
  - Otherwise the spatial continuity will be underestimated
- Should calculate the variogram on the variable being modeled with transforms (data and coordinates)
- Calculate the variogram, as this is what we model and apply in estimation and simulation (more later).

Choosing the Directions

- Inspect the data and interpretations, sections, plan views, …
- Azimuth angles in degrees clockwise from north
- Review multiple directions before choosing a set of 3 perpendicular directions
  – Omnidirectional: all directions taken together → often yields the most well-behaved variograms.
  – Major horizontal direction & two perpendicular to major direction
  – All anisotropy in geostatistics is geometric – three mutually orthogonal directions with ellipsoidal change in the other directions:

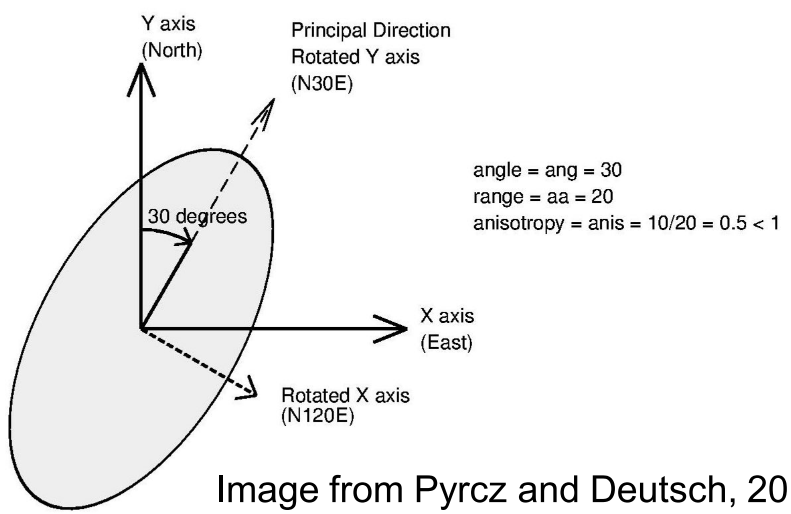


Image from Pyrcz and Deutsch, 2014

Choosing the Lag Distances and Tolerances

Guidance for Variogram Calculation Parameters:

- Lag separation distance should coincide with data spacing
- Lag tolerance typically chosen to be ½ lag separation distance
  - in cases of erratic variograms, may choose to overlap calculations so lag tolerance > ½ lag separation, results in more pairs.
- The variogram is only valid for a distance one half of the field size: start leaving data out of calculations with larger distances

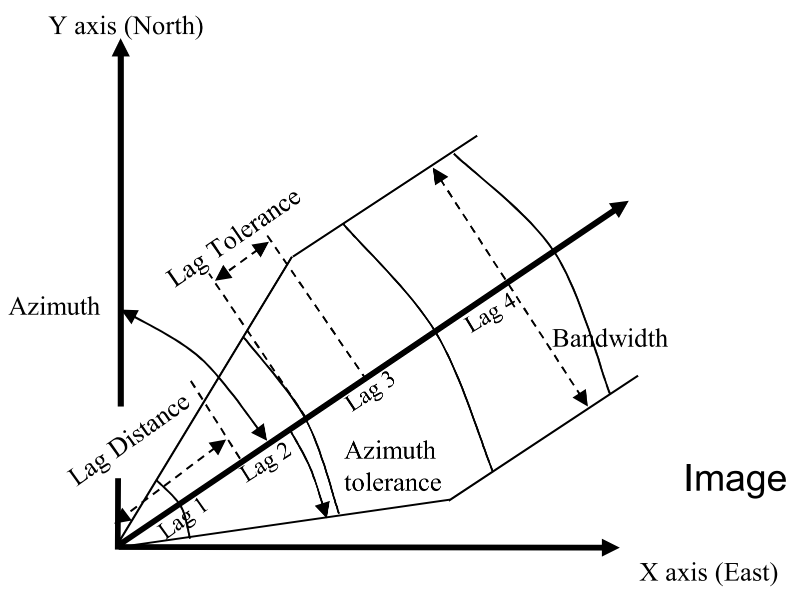
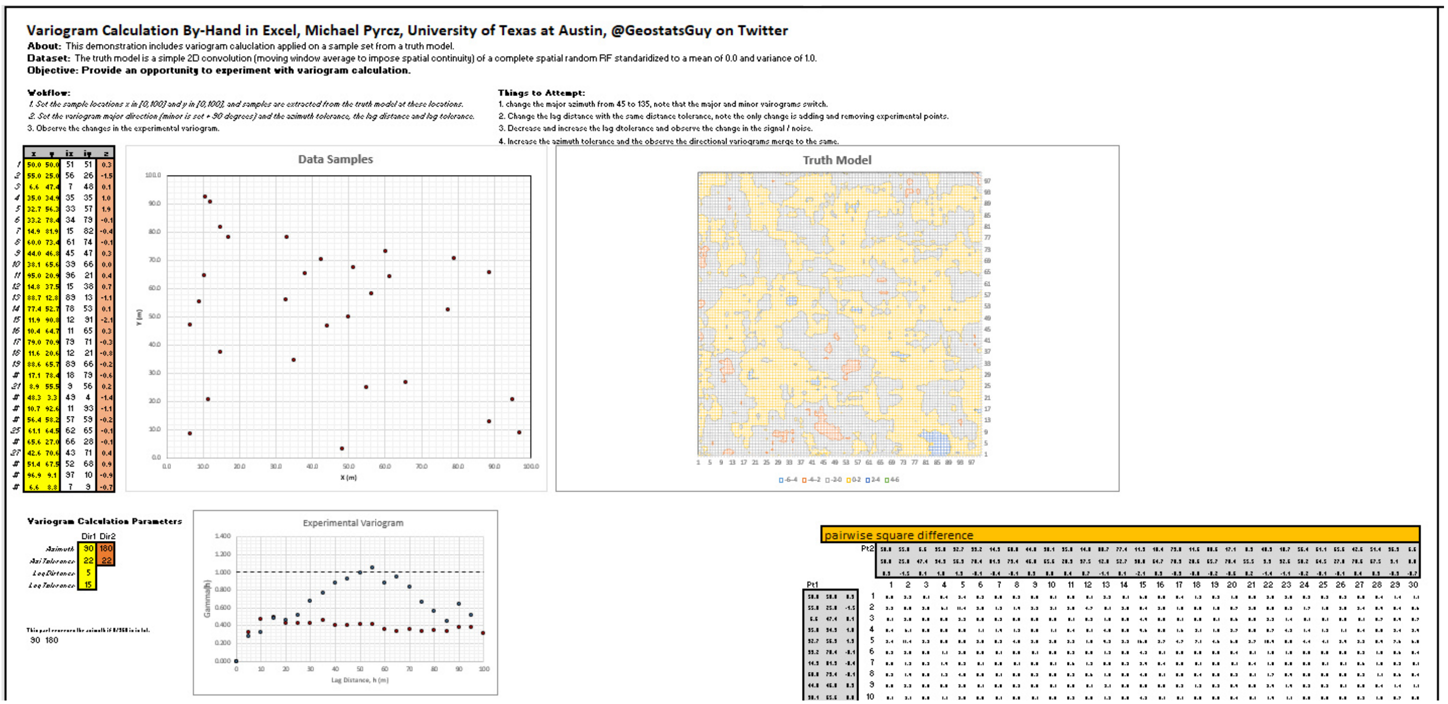


Image from Pyrcz and Deutsch, 2014

Spatial Calculation in Hands-on in Excel

Experiment with Variogram Calculation:



Things to try:

1. Set azimuth / lag tolerance small and large
2. Change directions.

The file is Variogram_Calc_Model_Demo_v2.0.xlsx.

Spatial Calculation in Demo in Python

Experiment with Variogram Calculation:

Things to Try:

Variogram maps

- Relate to the data location maps

Directional variograms

- Change lag tolerance
- Change lag distance

GeostatsPy: Spatial Continuity Directions 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: Methods to Detect Directions of Continuity with GeostatsPy

Here's a simple workflow on detecting the major spatial continuity directions in a spatial dataset with variogram analysis. This information is essential to optimum well placement and prectiction away from wells. First let's explain the concept of spatial continuity and the variogram.

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.

- The expected (average) squared difference between values separated by a lag distance vector (distance and direction), h:

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

where $z(\mathbf{u}_\alpha)$ and $z(\mathbf{u}_\alpha + \mathbf{h})$ are the spatial sample values at tail and head locations of the lag vector respectively.

- Calculated over a suite of lag distances to obtain a continuous function.
- the $\frac{1}{2}$ term converts a variogram into a semivariogram, but in practice the term variogram is used instead of semivariogram.
- We prefer the semivariogram because it relates directly to the covariance function, $C_x(\mathbf{h})$ and univariate variance, σ_x^2:

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

Note the correlogram is related to the covariance function as:

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

The correlogram provides of function of the $\mathbf{h} - \mathbf{h}$ scatter plot correlation vs. lag offset $\mathbf{h}$.

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

Variogram Observations

The following are common observations for variograms that should assist with their practical use.

Observation #1 - As distance increases, variability increase (in general).

Python notebook file: GeostatsPy_spatial_continuity_directions.ipynb

Spatial Continuity New Tools

Topic	Application to Subsurface Modeling
Stationarity	In the presence of multivariate relationships, must jointly model variables. *Summarize with bivariate statistics, and visualize and use conditional statistics to go beyond linear measures.*
Random Variables and Functions	Random variables and random functions are used to represent spatial uncertainty. *Porosity at a pre-drill location has the uncertainty model based on a random variable with Gaussian mean of 15% and standard deviation of 3%.*
Variogram Calculation	*Calculate spatial continuity from spatial data to use for spatial prediction.* *From the available wells the porosity spatial continuity range is 300 m in the 030 azimuth, we have no information beyond this spacing from existing wells.*

Geostatistics and Machine Learning
Spatial Continuity and Prediction

Lecture outline . . .

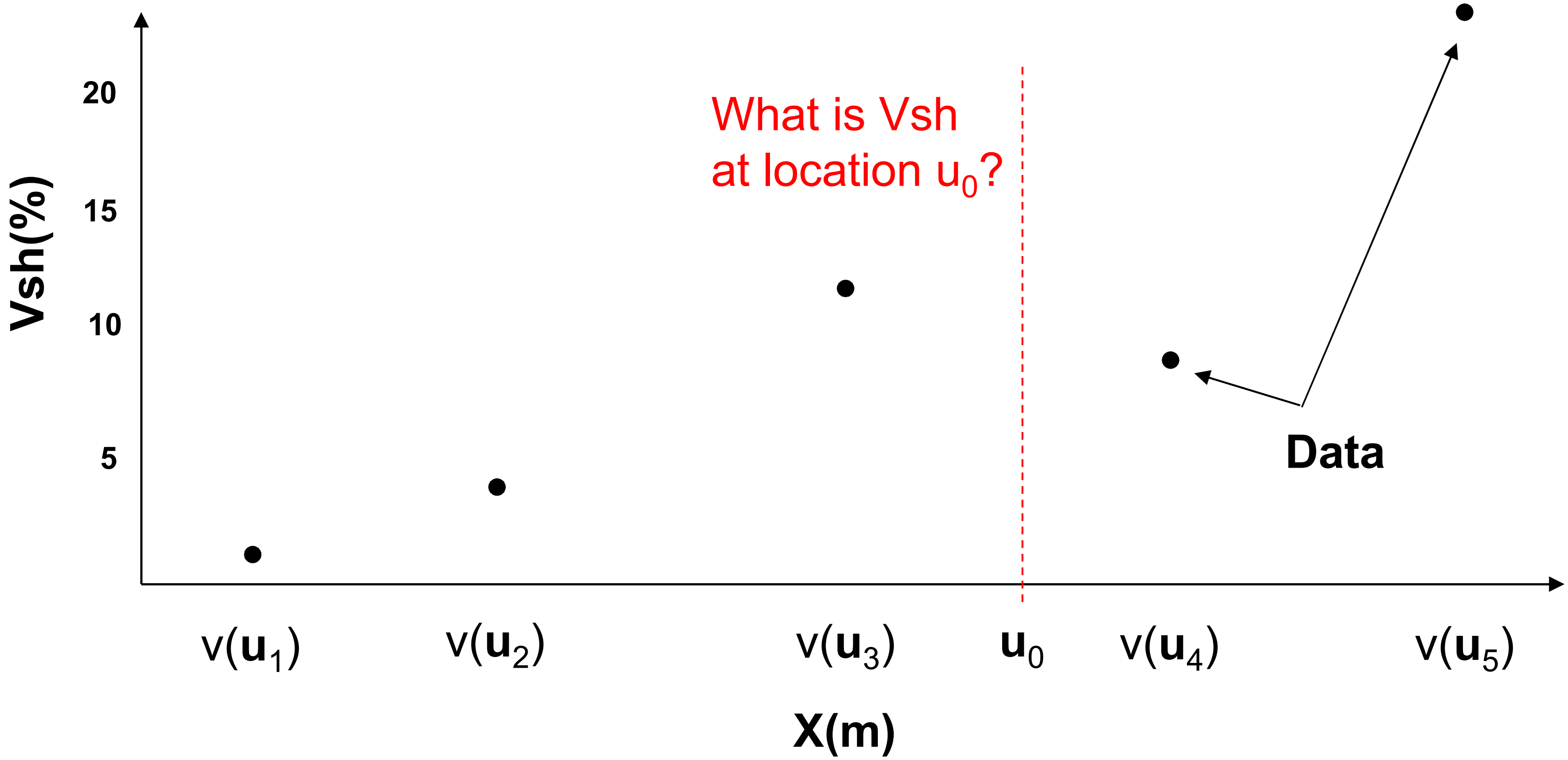
- **Spatial Estimation**

Instructor: Michael Pyrcz, the University of Texas at Austin

Trend Modeling

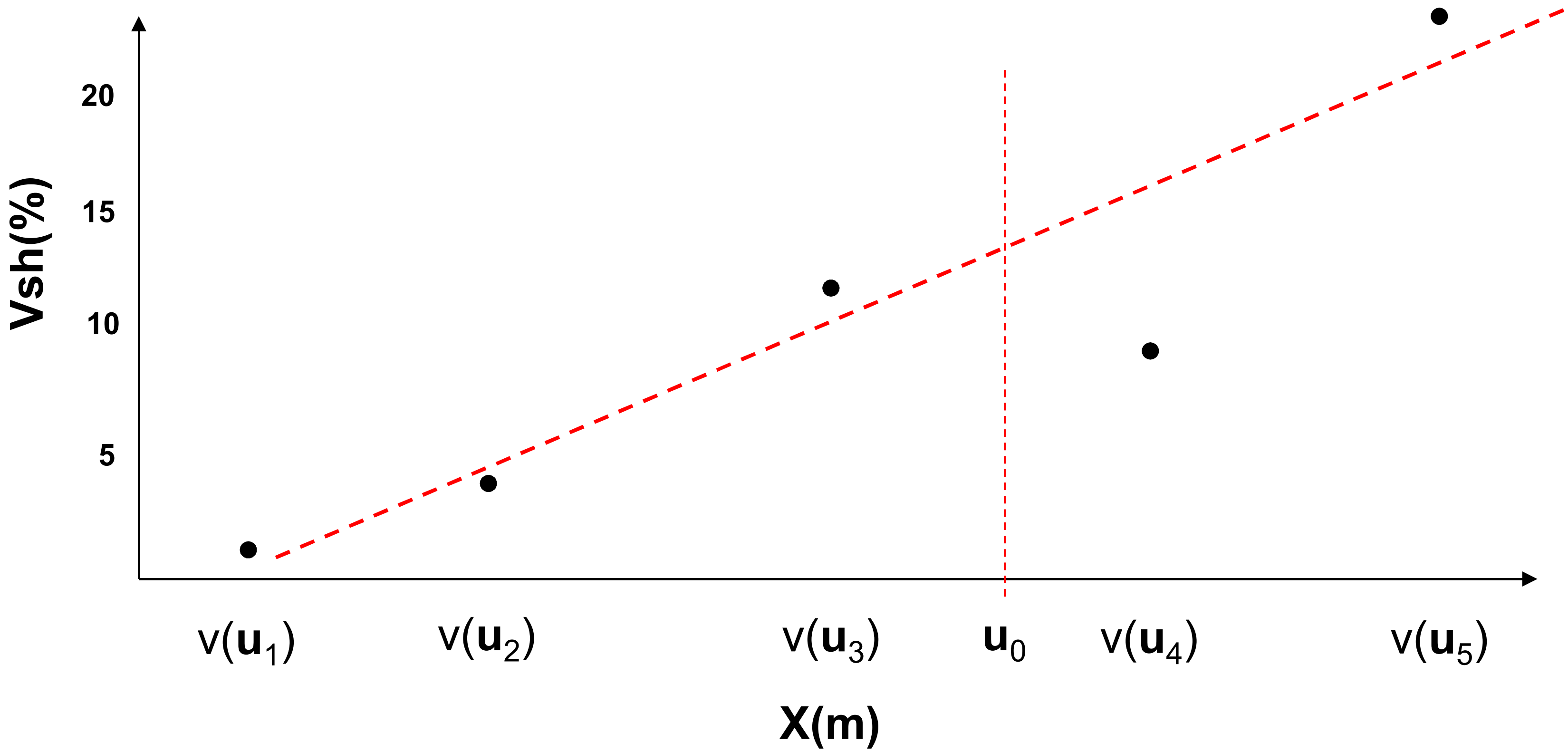
We Must Start with Trend Modeling

- Geostatistical spatial estimation methods make an assumption concerning stationarity
  - In the presence of significant nonstationarity we would not rely 100% for spatial estimation on data + spatial continuity model



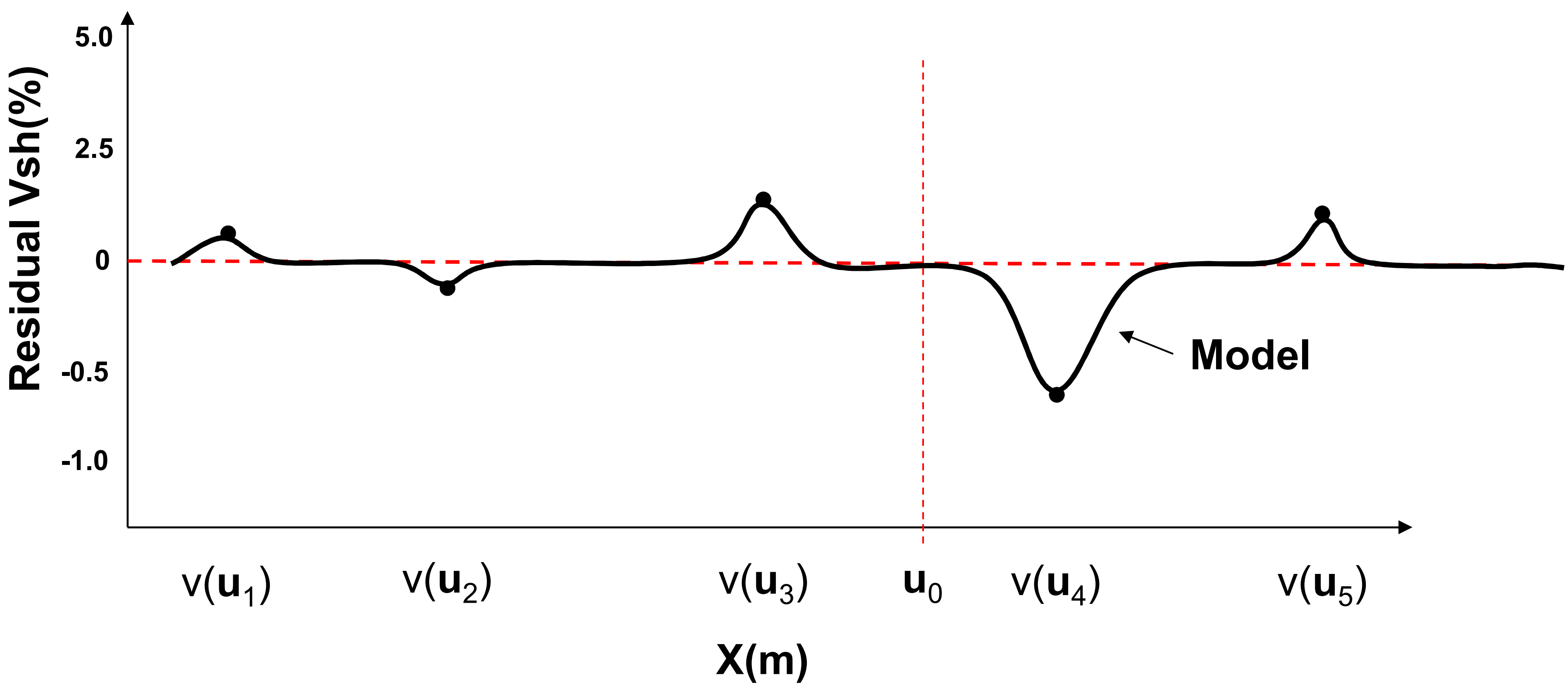
Trend Modeling

- Geostatistical spatial estimation methods will make an assumption concerning stationarity
  - If we observe a trend, we should model the trend.



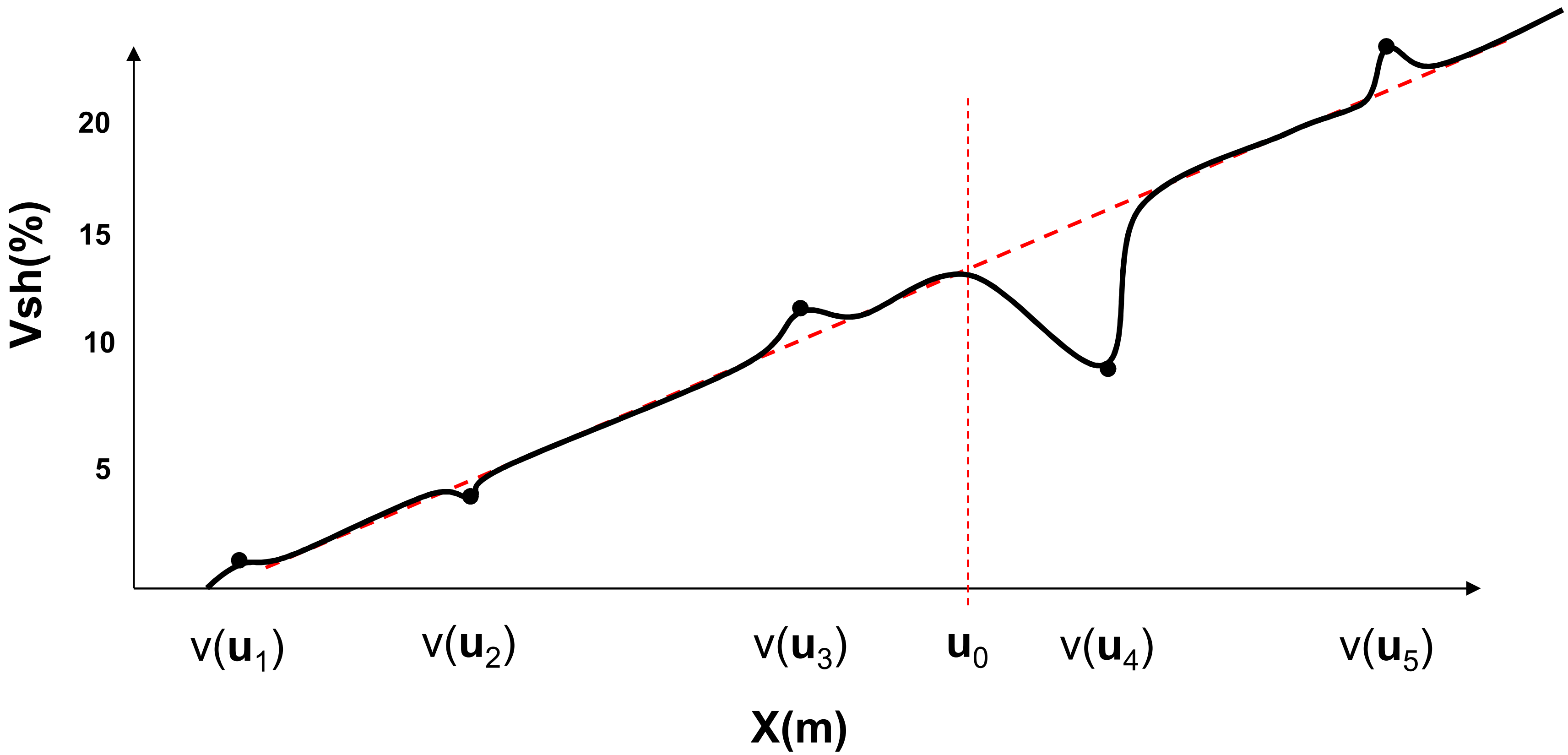
Trend Modeling

- Geostatistical spatial estimation methods will make an assumption concerning stationarity
  - Then model the residuals.



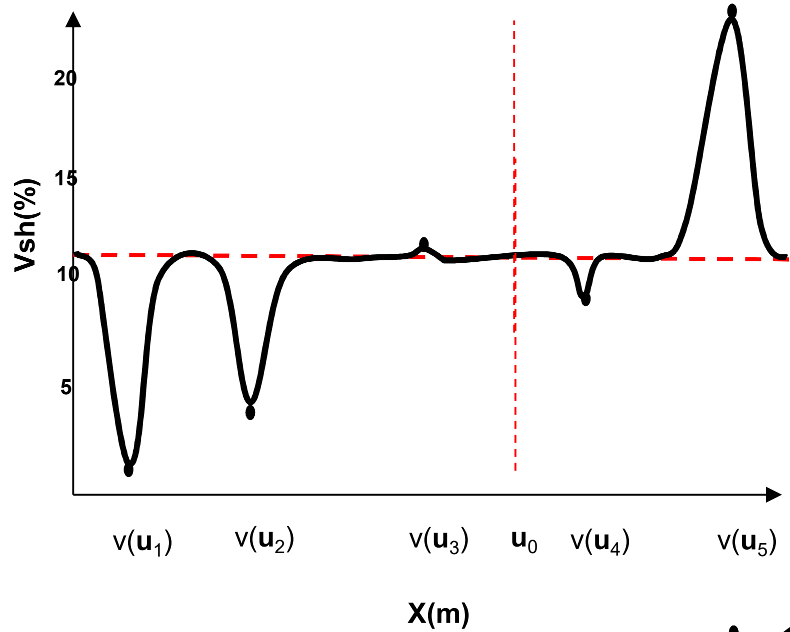
Trend Modeling

- Geostatistical spatial estimation methods will make an assumption concerning stationarity
  - After modeling, add the trend back to the modelled residuals

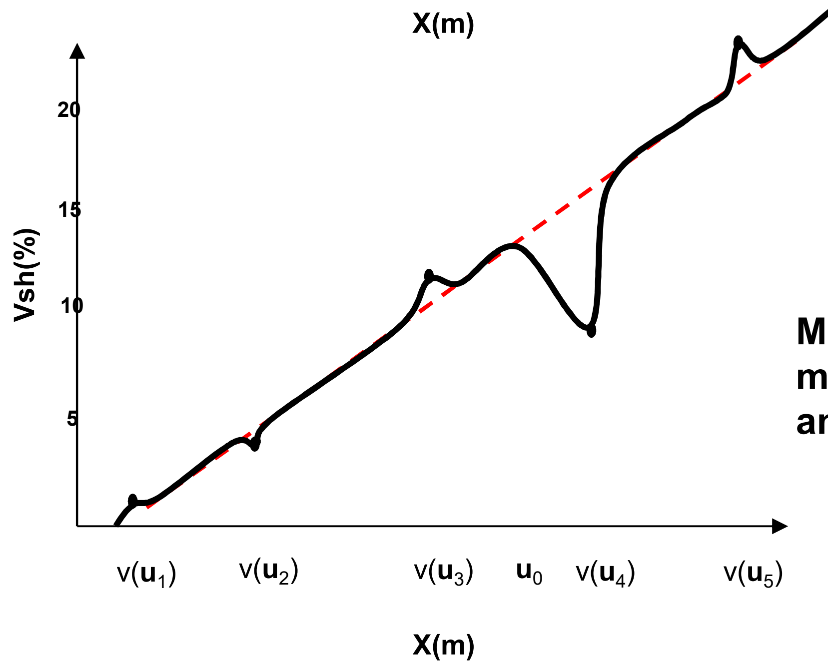


Trend Modeling

– How bad could it be if we did not model a trend?
– Geostatistical estimation would assume stationarity and away from data we would estimate with the global mean (simple kriging)!



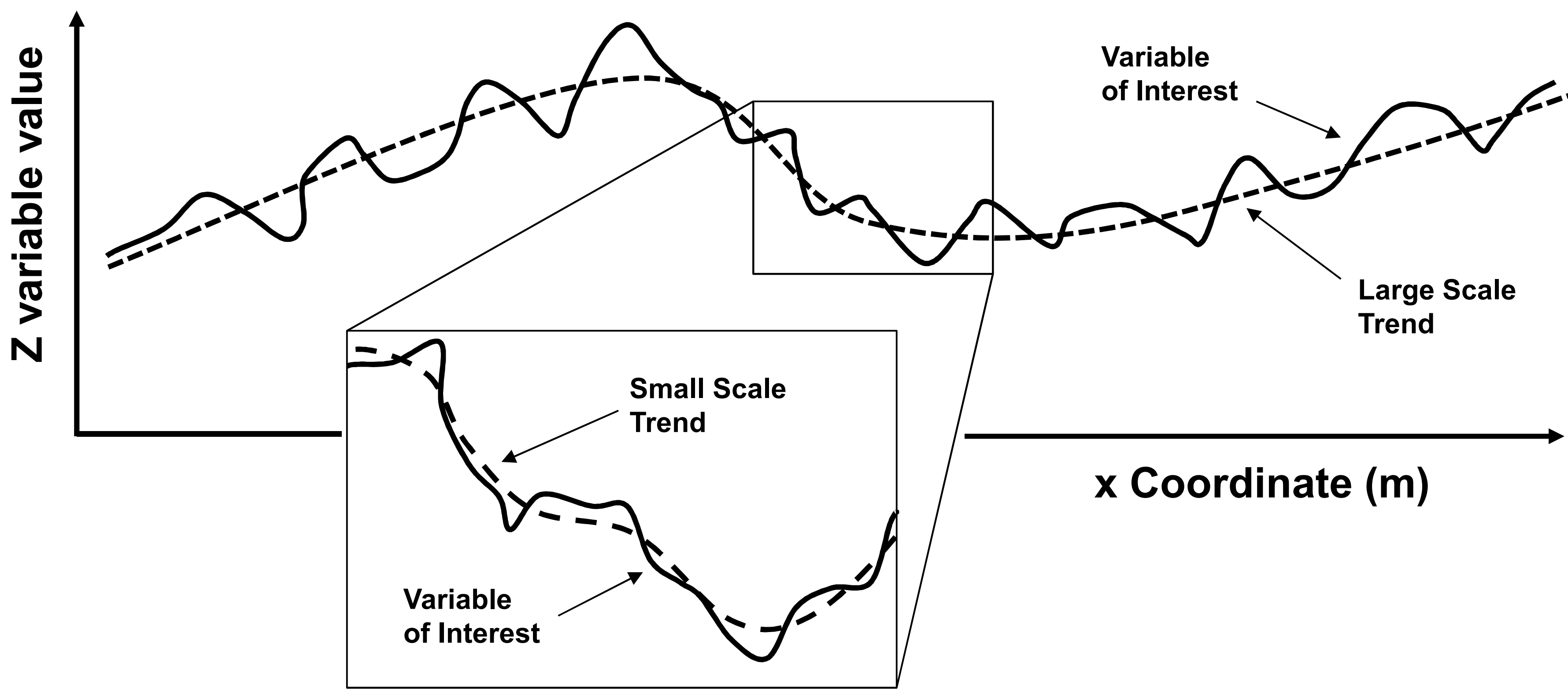
Model with stationary mean + data.



Model with mean trend model and residual + data.

Trend Modeling

- Trend Modeling
  - We must identify and model trends / nonstationarities



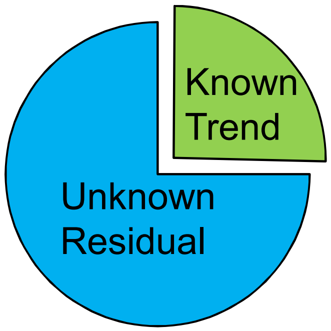
Trend Modeling

- Any variance in the trend is removed from the residual:

$$\boldsymbol{\sigma_X^2 = \sigma_{X_t}^2 + \sigma_{X_r}^2 + 2C_{X_t,X_r}}$$

  - if the $\boldsymbol{X_t \perp\!\!\!\perp X_r}$, $\boldsymbol{C_{X_t,X_r}} = 0$

$$\boldsymbol{\sigma_{X_r}^2 = \sigma_X^2 - \sigma_{X_t}^2}$$



  – So if $\boldsymbol{\sigma_X^2}$ is the total variance (variability), and $\boldsymbol{\sigma_{X_t}^2}$ is the variability that is deterministically modelled, treated as known, and $\boldsymbol{\sigma_{X_r}^2}$ is the component of the variability that is treated as unknown.
  – Result: the more variability explained by the trend the less variability that remains as uncertain.

Definition Deterministic Model

- Model that assumes perfect knowledge, without uncertainty

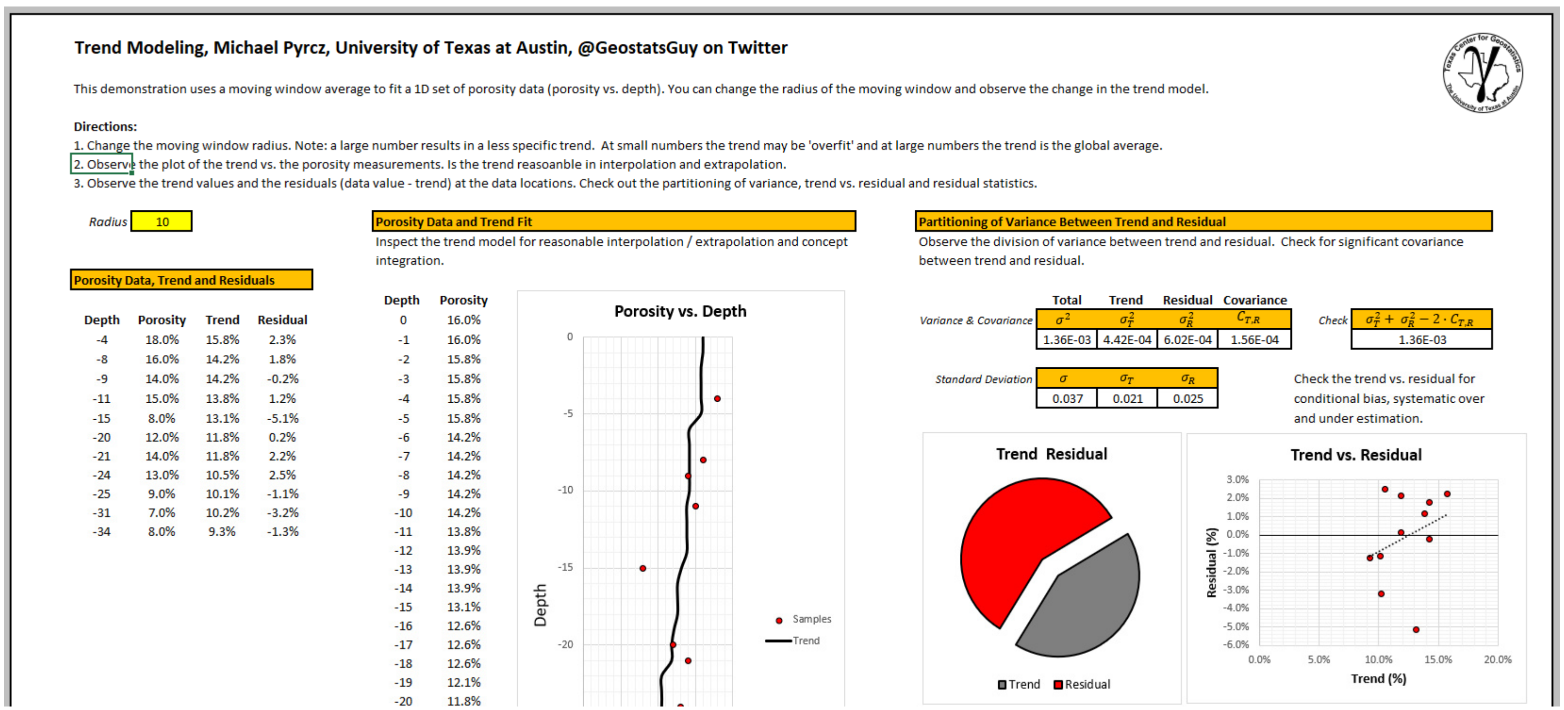
- Based on knowledge of the phenomenon or trend fitting to data
- Most subsurface models have a deterministic component (trend) to capture expert knowledge and to provide a stationary residual for geostatistical modeling.

Trend Modeling Hands-on

Here's an opportunity for experiential learning with Trend Modeling.

- **Things to try:**

1. Set the radius very large (50). How's the trend model performing? Try radius very small (1).
2. What do you think is the best radius to fit a trend to this spatial data?



Trend Modeling, Michael Pyrcz, University of Texas at Austin, @GeostatsGuy on Twitter

This demonstration uses a moving window average to fit a 1D set of porosity data (porosity vs. depth). You can change the radius of the moving window and observe the change in the trend model.

Directions:
1. Change the moving window radius. Note: a large number results in a less specific trend. At small numbers the trend may be 'overfit' and at large numbers the trend is the global average.
2. Observe the plot of the trend vs. the porosity measurements. Is the trend reasoanble in interpolation and extrapolation.
3. Observe the trend values and the residuals (data value - trend) at the data locations. Check out the partitioning of variance, trend vs. residual and residual statistics.

Radius 10

Porosity Data, Trend and Residuals

Depth	Porosity	Trend	Residual
-4	18.0%	15.8%	2.3%
-8	16.0%	14.2%	1.8%
-9	14.0%	14.2%	-0.2%
-11	15.0%	13.8%	1.2%
-15	8.0%	13.1%	-5.1%
-20	12.0%	11.8%	0.2%
-21	14.0%	11.8%	2.2%
-24	13.0%	10.5%	2.5%
-25	9.0%	10.1%	-1.1%
-31	7.0%	10.2%	-3.2%
-34	8.0%	9.3%	-1.3%

Porosity Data and Trend Fit

Inspect the trend model for reasonable interpolation / extrapolation and concept integration.

Depth	Porosity
0	16.0%
-1	16.0%
-2	15.8%
-3	15.8%
-4	15.8%
-5	15.8%
-6	14.2%
-7	14.2%
-8	14.2%
-9	14.2%
-10	14.2%
-11	13.8%
-12	13.9%
-13	13.9%
-14	13.9%
-15	13.1%
-16	12.6%
-17	12.6%
-18	12.6%
-19	12.1%
-20	11.8%

Partitioning of Variance Between Trend and Residual

Observe the division of variance between trend and residual. Check for significant covariance between trend and residual.

	Total	Trend	Residual	Covariance
Variance & Covariance	σ^2	σ_T^2	σ_R^2	$C_{T,R}$
	1.36E-03	4.42E-04	6.02E-04	1.56E-04

Check $\sigma_T^2 + \sigma_R^2 - 2 \cdot C_{T,R}$ = 1.36E-03

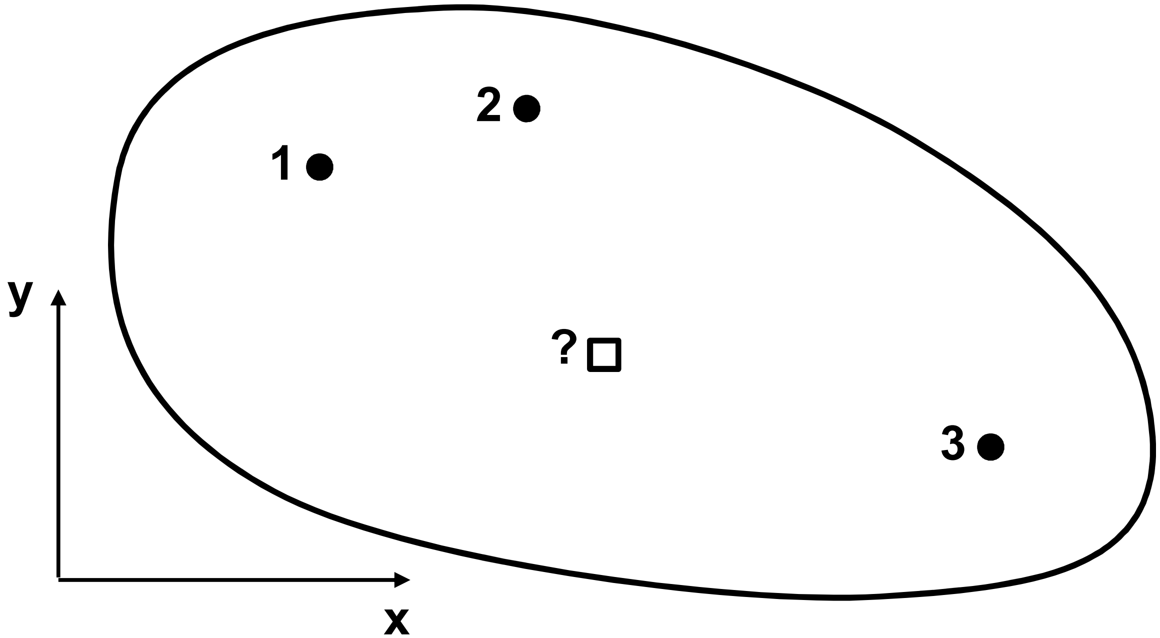
Standard Deviation	σ	σ_T	σ_R
	0.037	0.021	0.025

Check the trend vs. residual for conditional bias, systematic over and under estimation.

File Name: Trend_Modeling.xlsx File is at: https://git.io/fhALP

Spatial Estimation

- Consider the case of estimating at some unsampled location:

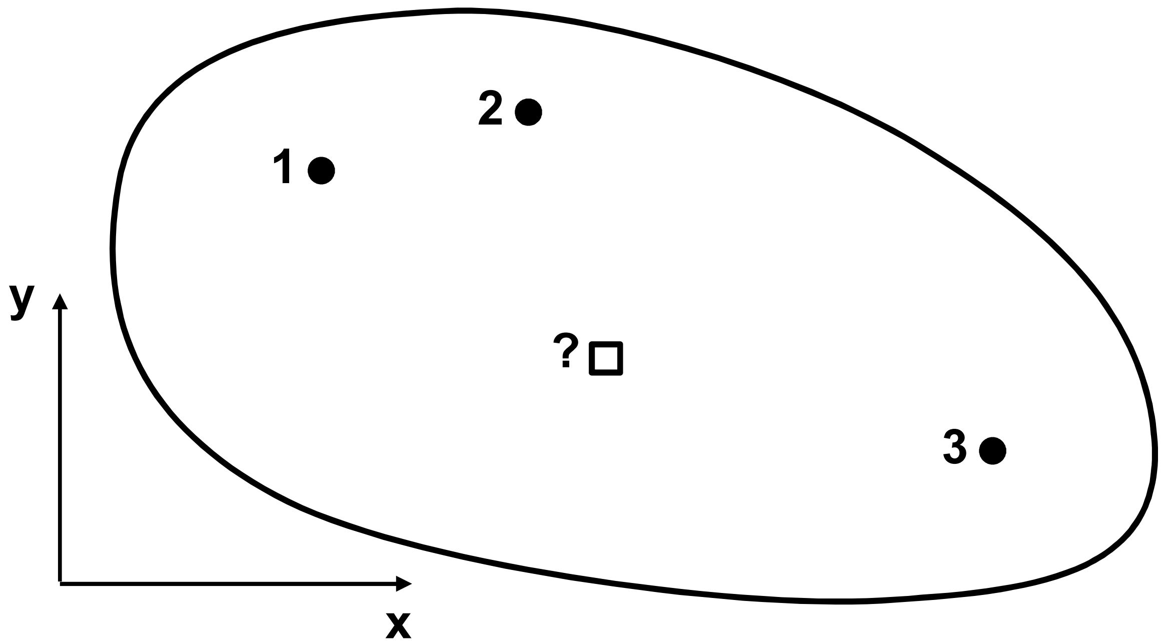


- How would you do this given data, $z(\mathbf{u}_1)$, $z(\mathbf{u}_2)$, and $z(\mathbf{u}_3)$?

- Note: z is the variable of interest (e.g. porosity etc.) and $\mathbf{u}_i$ is the data locations.

Spatial Estimation

- Consider the case of estimating at some unsampled location:



$\mathbf{z}(\mathbf{u}_\alpha)$ **is the data values**

$z^*(\mathbf{u}_0)$ **is an estimate**

λ_α **is the data weights**

m_z **is the global mean**

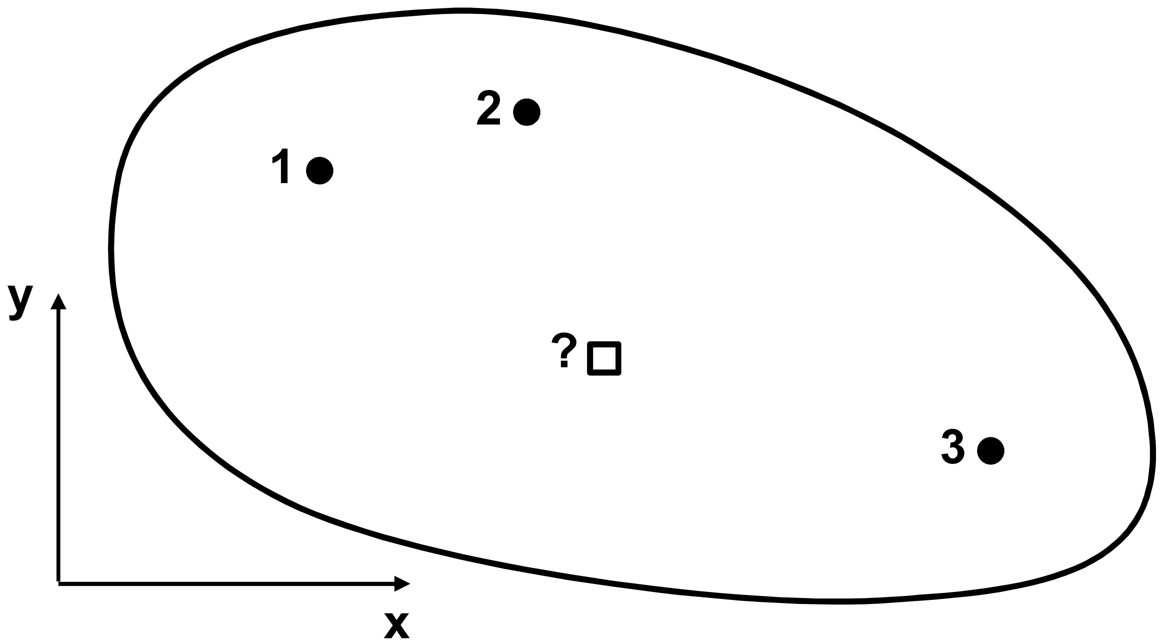
- How would you do this given data, $z(\mathbf{u}_1)$, $z(\mathbf{u}_2)$, and $z(\mathbf{u}_3)$?

$$z^*(\mathbf{u}_0) = \sum_{\alpha=1}^{n} \lambda_\alpha \mathbf{z}(\mathbf{u}_\alpha) + \boxed{\left(1 - \sum_{\alpha=1}^{n} \lambda_\alpha\right) m_z}$$

Unbiasedness Constraint Weights sum to 1.0.

Spatial Estimation

- Consider the case of estimating at some unsampled location:



- How would you do this given data, $z(\mathbf{u}_1)$, $z(\mathbf{u}_2)$, and $z(\mathbf{u}_3)$?

$$z^*(\mathbf{u}_0) - m_z(\mathbf{u}_0) = \sum_{\alpha=1}^{n} \lambda_\alpha \big(\mathrm{z}(\mathbf{u}_\alpha) - m_z(\mathbf{u}_\alpha)\big)$$

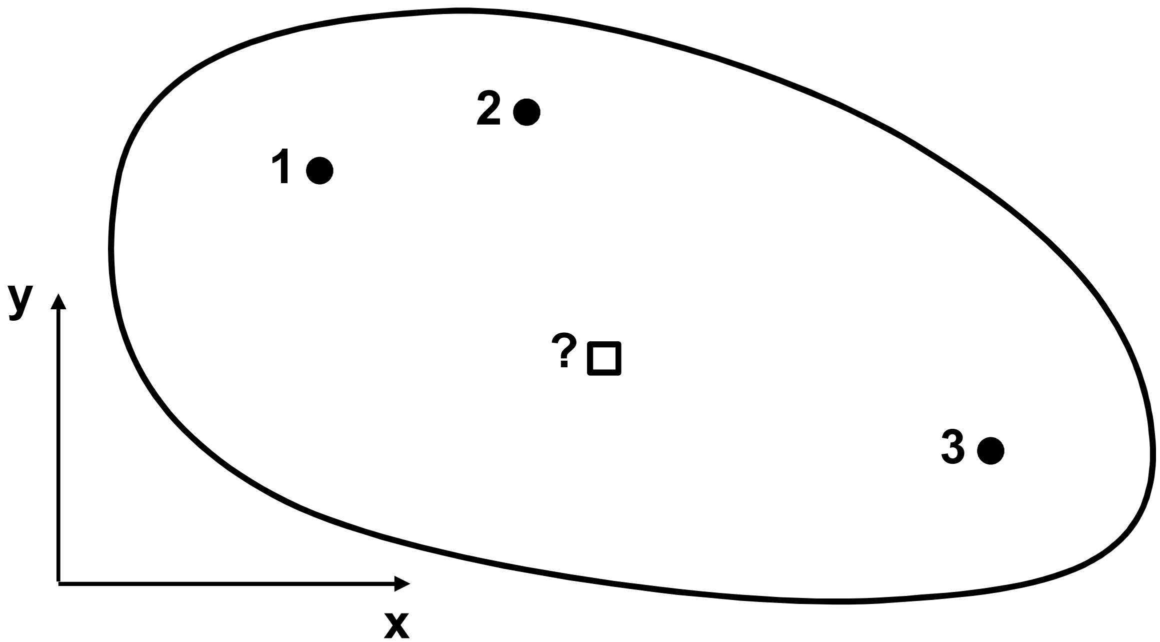
In the case where the mean is non-stationary.

Given y$= z - m, \quad y^*(\mathbf{u}_0) = \sum_{\alpha=1}^{n} \lambda_\alpha \, \mathrm{y}(\mathbf{u}_\alpha)$

Simplified with residual, y.

Spatial Estimation

- Consider the case of estimating at some unsampled location:



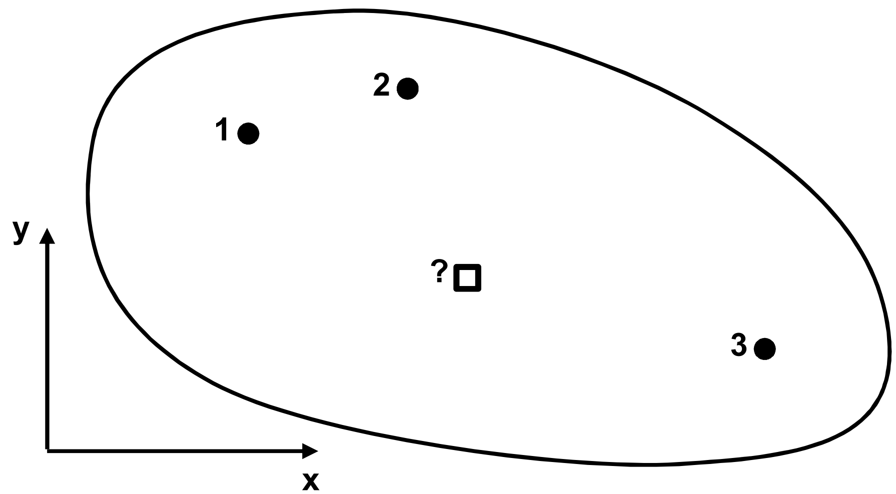
- Linear weighted, sound good. How do we get the weights? $\boldsymbol{\lambda_\alpha, \alpha = 1, \ldots, n}$

$$y^*(\mathbf{u_0}) = \sum_{\alpha=1}^{n} \lambda_\alpha \, \mathrm{y}(\mathbf{u}_\alpha)$$

Simplified with residual, y.

Spatial Estimation

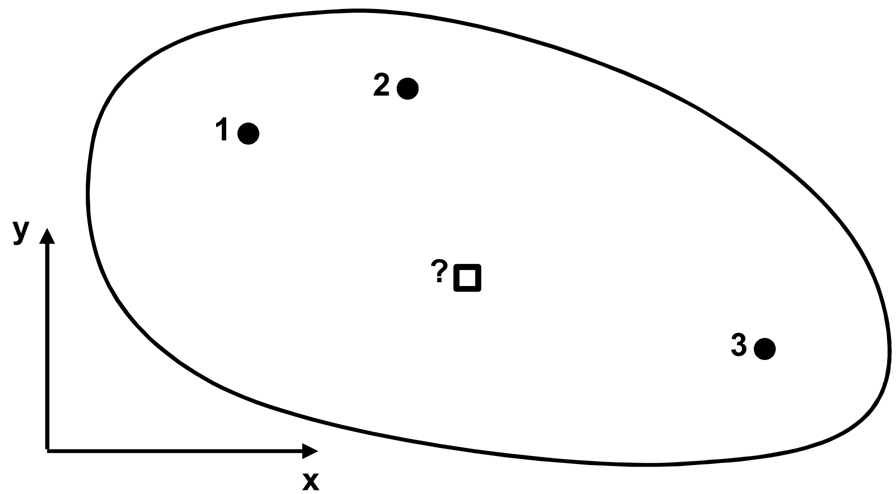
- Consider the case of estimating at some unsampled location:



- Linear weighted, sound good. How do we get the weights? $\boldsymbol{\lambda_{\alpha}, \alpha = 1, \ldots, n}$
- Equal weighted / average? $\boldsymbol{\lambda_{\alpha} = {}^{1}/_{n}}$ **Equal weight average of data**
- What's wrong with that?

Spatial Estimation

- Consider the case of estimating at some unsampled location:



- How do we get the weights? $\boldsymbol{\lambda_\alpha, \alpha = 1, \ldots, n}$

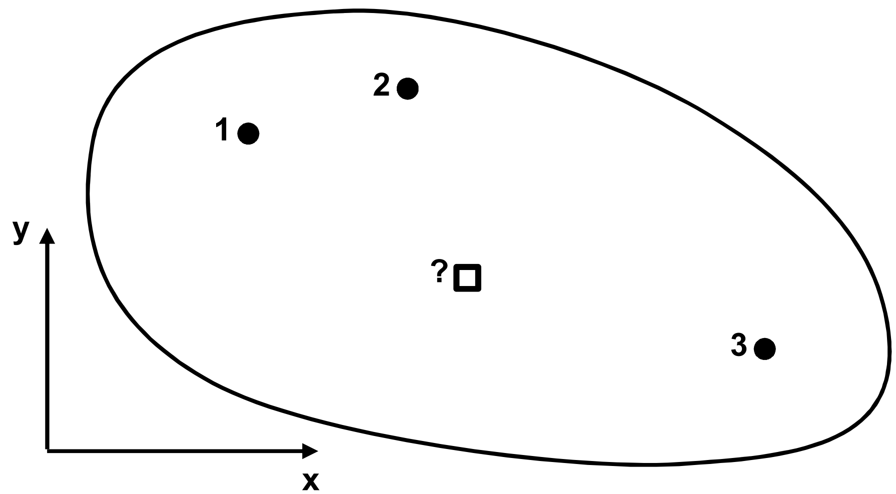
Inverse distance to power standardized so weights sum to 1.0.

- Inverse distance? $$\lambda_\alpha = \frac{1}{dist(\mathbf{u}_0, \mathbf{u}_\alpha)^p} \Big/ \sum_{\alpha=1}^{n} \lambda_\alpha$$

- What's wrong with that?

Spatial Estimation

- Consider the case of estimating at some unsampled location:



- How do we get the weights? $\boldsymbol{\lambda_\alpha, \alpha = 1, \ldots, n}$

- It would be great to use weight that account for closeness (spatial correlation > distance alone), redundancy (once again with spatial correlation).

- How can we do that?

Spatial Estimation

Kriging Derivations Removed

The results is a very useful and interpretable linear system of equations.

Simple Kriging: Some Details

There are three equations to determine the three weights:

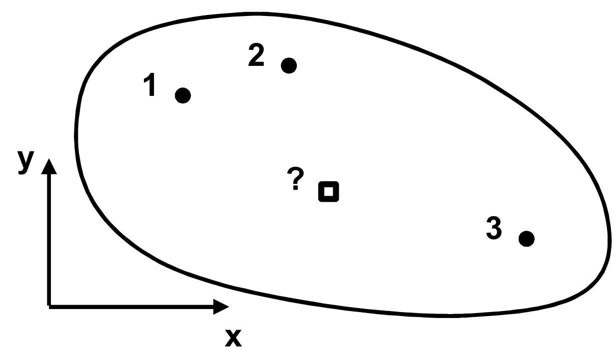
$$\lambda_1 \cdot C(\mathbf{u}_1, \mathbf{u}_1) + \lambda_2 \cdot C(\mathbf{u}_1, \mathbf{u}_2) + \lambda_3 \cdot C(\mathbf{u}_1, \mathbf{u}_3) = C(\mathbf{u}_0, \mathbf{u}_1)$$

$$\lambda_1 \cdot C(\mathbf{u}_2, \mathbf{u}_1) + \lambda_2 \cdot C(\mathbf{u}_2, \mathbf{u}_2) + \lambda_3 \cdot C(\mathbf{u}_2, \mathbf{u}_3) = C(\mathbf{u}_0, \mathbf{u}_2)$$

$$\lambda_1 \cdot C(\mathbf{u}_3, \mathbf{u}_1) + \lambda_2 \cdot C(\mathbf{u}_1, \mathbf{u}_2) + \lambda_3 \cdot C(\mathbf{u}_1, \mathbf{u}_3) = C(\mathbf{u}_0, \mathbf{u}_1)$$

In matrix notation: Recall that $C(\mathbf{h}) = C(0) - \gamma(\mathbf{h})$

$$\underbrace{\begin{bmatrix} C(\mathbf{u}_1, \mathbf{u}_1) & C(\mathbf{u}_1, \mathbf{u}_2) & C(\mathbf{u}_1, \mathbf{u}_3) \\ C(\mathbf{u}_2, \mathbf{u}_1) & C(\mathbf{u}_2, \mathbf{u}_2) & C(\mathbf{u}_2, \mathbf{u}_3) \\ C(\mathbf{u}_3, \mathbf{u}_1) & C(\mathbf{u}_3, \mathbf{u}_2) & C(\mathbf{u}_3, \mathbf{u}_3) \end{bmatrix}}_{\text{redundancy}} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix}}_{\text{weights}} = \underbrace{\begin{bmatrix} C(\mathbf{u}_o, \mathbf{u}_1) \\ C(\mathbf{u}_o, \mathbf{u}_2) \\ C(\mathbf{u}_o, \mathbf{u}_3) \end{bmatrix}}_{\text{closeness}}$$



Properties of Simple Kriging

What Does Kriging Provide?:

- **Best Estimate**: Minimum error estimator (just try to pick weights, you won't bet it)

- **Estimation Variance**: Provides a measure of the estimation (or kriging) variance (uncertainty in the estimate):

$$\sigma_E^2(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u} - \mathbf{u}_\alpha) \qquad \boldsymbol{\sigma_E^2} \rightarrow [\mathbf{0}, \boldsymbol{\sigma_x^2}]$$

Estimation Variance

Variance

Kriging Weights

Covariance Between Data and Estimate Location

More Properties

- Exact interpolator: at data location
- Kriging variance can be calculated before getting the sample information, homoscedastic!
- Kriging takes into account:
  - distance of the information: $C(\mathbf{u},\mathbf{u}_i)$
  - configuration of the data: $C(\mathbf{u}_i,\mathbf{u}_j)$
  - structural continuity of the variable being considered: $C(\mathbf{h})$
- The smoothing effect of kriging can be forecast – we will return to this with simulation.
- Kriging theory is part of the probabilistic theory of *projectors*: orthogonal projection onto space of linear combinations of the *n* data (Hilbert space)

Simple Kriging Hands-on

Here's an opportunity for experiential learning with Simple Kriging.

- **Things to try:**

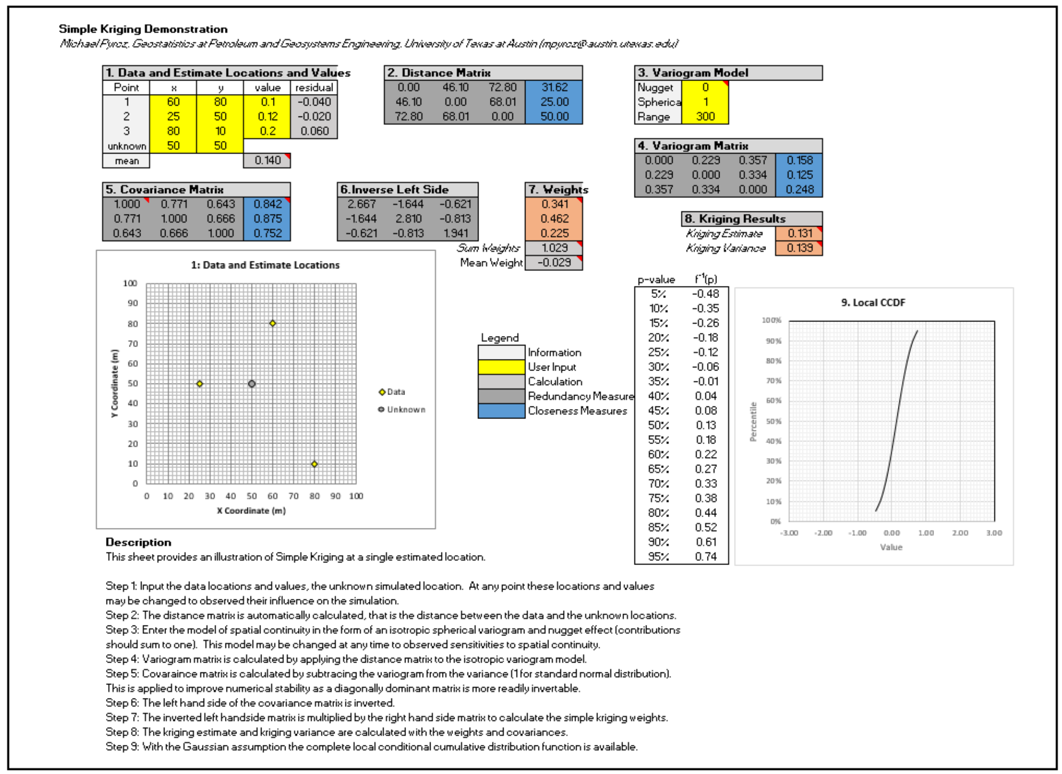
Pay attention to the kriging weights, kriging estimate and kriging variance while you:

1. Set points 1 and 2 closer together.

2. Put point 1 behind point 2 to create screening.

3. Put all points outside the range.

4. Set the range very large.



File Name: Simple_Kriging_Demo.xlsx File is at: https://git.io/fNgBK

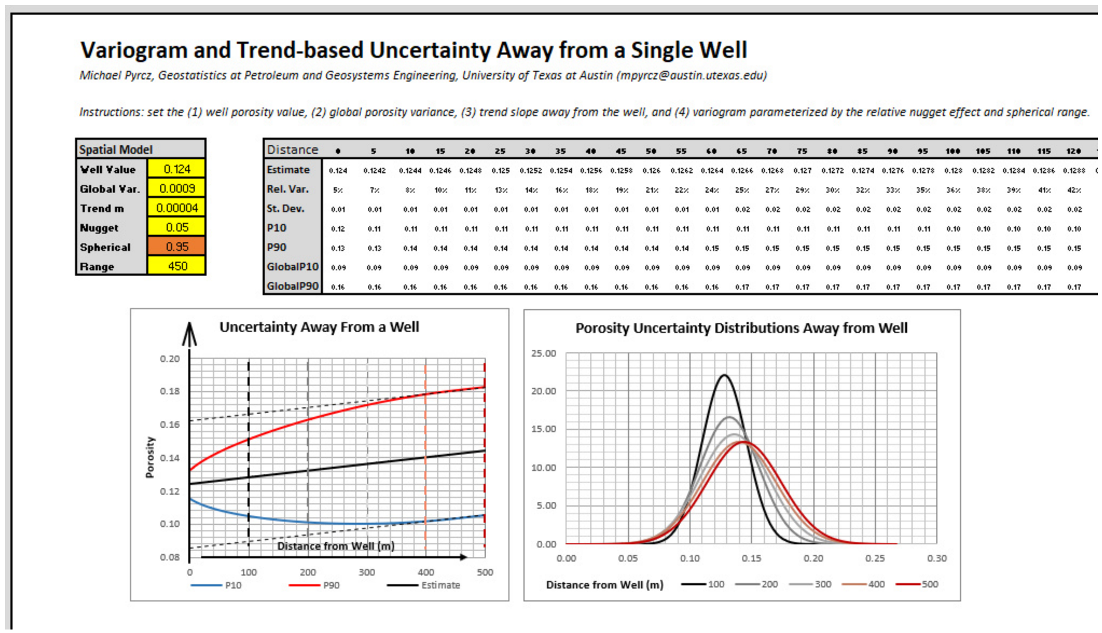
Spatial Uncertainty Hands-on

Here's an opportunity for experiential learning with Simple Kriging for spatial uncertainty. The kriging estimation variance is very useful.

- **Things to try:**

Pay attention to the kriging uncertainty P10, mean and P90 away from the well as you:

1. Change the spatial continuity range.
2. Add and adjust the nugget effect.
3. Modify the trend slope.



Variogram and Trend-based Uncertainty Away from a Single Well

Michael Pyrcz, Geostatistics at Petroleum and Geosystems Engineering, University of Texas at Austin (mpyrcz@austin.utexas.edu)

Instructions: set the (1) well porosity value, (2) global porosity variance, (3) trend slope away from the well, and (4) variogram parameterized by the relative nugget effect and spherical range.

Spatial Model	
Well Value	0.124
Global Var.	0.0009
Trend m	0.00004
Nugget	0.05
Spherical	0.95
Range	450

Distance	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120
Estimate	0.124	0.1242	0.1244	0.1246	0.1248	0.125	0.1252	0.1254	0.1256	0.1258	0.126	0.1262	0.1264	0.1266	0.1268	0.127	0.1272	0.1274	0.1276	0.1278	0.128	0.1282	0.1284	0.1286	0.1288
Rel. Var.	5%	7%	8%	10%	11%	13%	14%	16%	18%	19%	21%	22%	24%	25%	27%	29%	30%	32%	33%	35%	36%	38%	39%	41%	42%
St. Dev.	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
P10	0.12	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.10	0.10	0.10
P90	0.13	0.13	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
GlobalP10	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
GlobalP90	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17

File Name: Uncertainty_Away_from_Well_Demo.xlsx   File is at: https://git.io/fNO2I

Spatial Estimation New Tools

Topic	Application to Subsurface Modeling
Trend Modeling	Decompose variance into deterministic trend and stochastic residual. *30% of porosity variance is described by a linear depth trend and 70% is described by a 3D variogram model.*
Kriging Estimates and Kriging Variances	Kriging provides the best estimate and a measure of estimation variance. *Given a kriging estimate of 13% and kriging variance of 9% and the assumption of a Gaussian distribution we have a complete local distribution of uncertainty for pre-drill porosity.*

Geostatistics and Machine Learning
Spatial Continuity and Prediction

Lecture outline . . .

- **Stationarity**
- **Spatial Continuity**
- **Variogram Calculation**
- **Spatial Estimation**

Introduction

Data Analytics

Inferential Methods

Predictive Methods

Advanced Methods

Conclusions

Instructor: Michael Pyrcz, the University of Texas at Austin