2-Day Course – Spatial Modeling with Geostatistics

Prof. Michael J. Pyrcz, Ph.D., P.Eng.
Associate Professor

Hildebrand Department of Petroleum & Geosystems Engineering
University of Texas at Austin

Bureau of Economic Geology, Jackson School of Geosciences
University of Texas at Austin

"In two days, what a geoscientists needs to know about geostatistics, and workflows to get you started with applying geostatistics to impact your work."

Spatial Modeling with Geostatistics

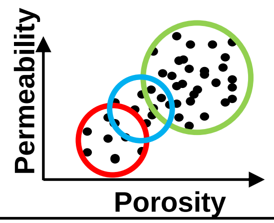
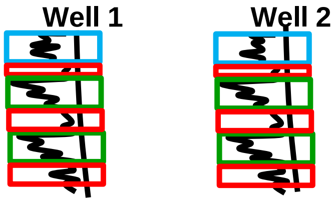
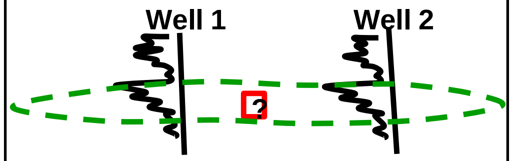
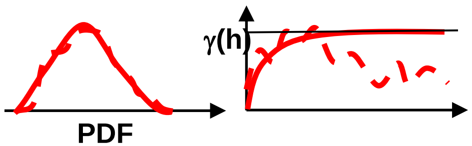
Spatial Data Analysis - Modeling

Lecture outline . . .

- **Interpreting Spatial Continuity**
- **Modeling Spatial Continuity**

Comments on Facies

- Facies are distinct, stationary parts of the subsurface (sandstone, shale)
- Facies must help with estimation away from measured locations. To do that they should meet criteria.

Criteria	Considerations	Example
Separation of Rock Properties	Facies must divide the properties of interest that impact subsurface environmental and economic performance (e.g. grade, porosity and permeability).	
Identifiable in Data	Facies must be identifiable with the most common data available. e.g. facies identifiable only in cores are not useful if most wells have only logs.	
Map-able Away from Data	Facies must be easier to predict away from data than the rock properties of interest directly, facies improves prediction.	
Sufficient Sampling	There must be enough data to allow for reliable inference of reliable statistics for rock properties for each facies.	

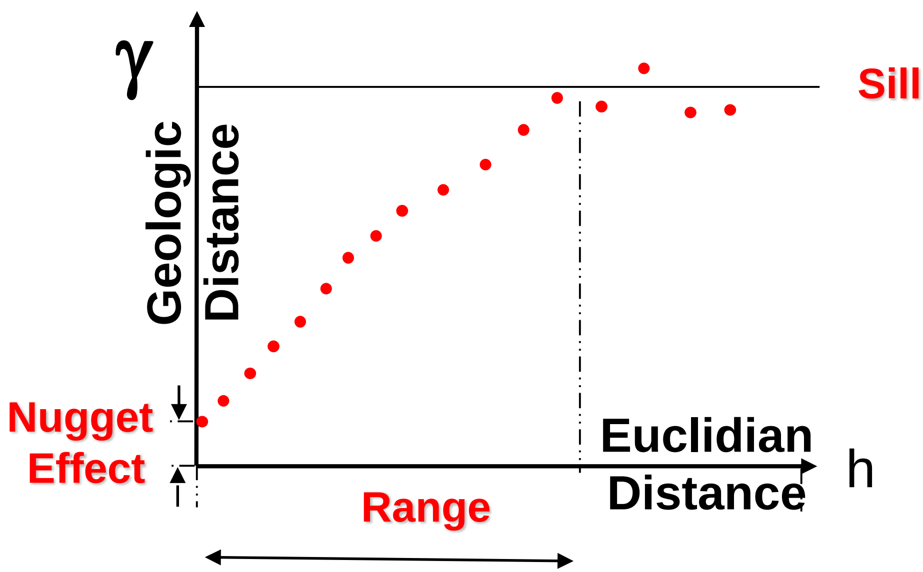
Variogram Terminology

- The variogram is a chart of *variance versus distance* or *geological variability versus direction and Euclidean distance.*
- The ***sill*** is the variance of the data used for variogram calculation (1.0 if the data are normal scores)
- The ***range*** is the distance at which the variogram reaches the sill
- The ***nugget effect*** is the behavior at distances less than the smallest experimental lag:

geological microstructure + measurement error

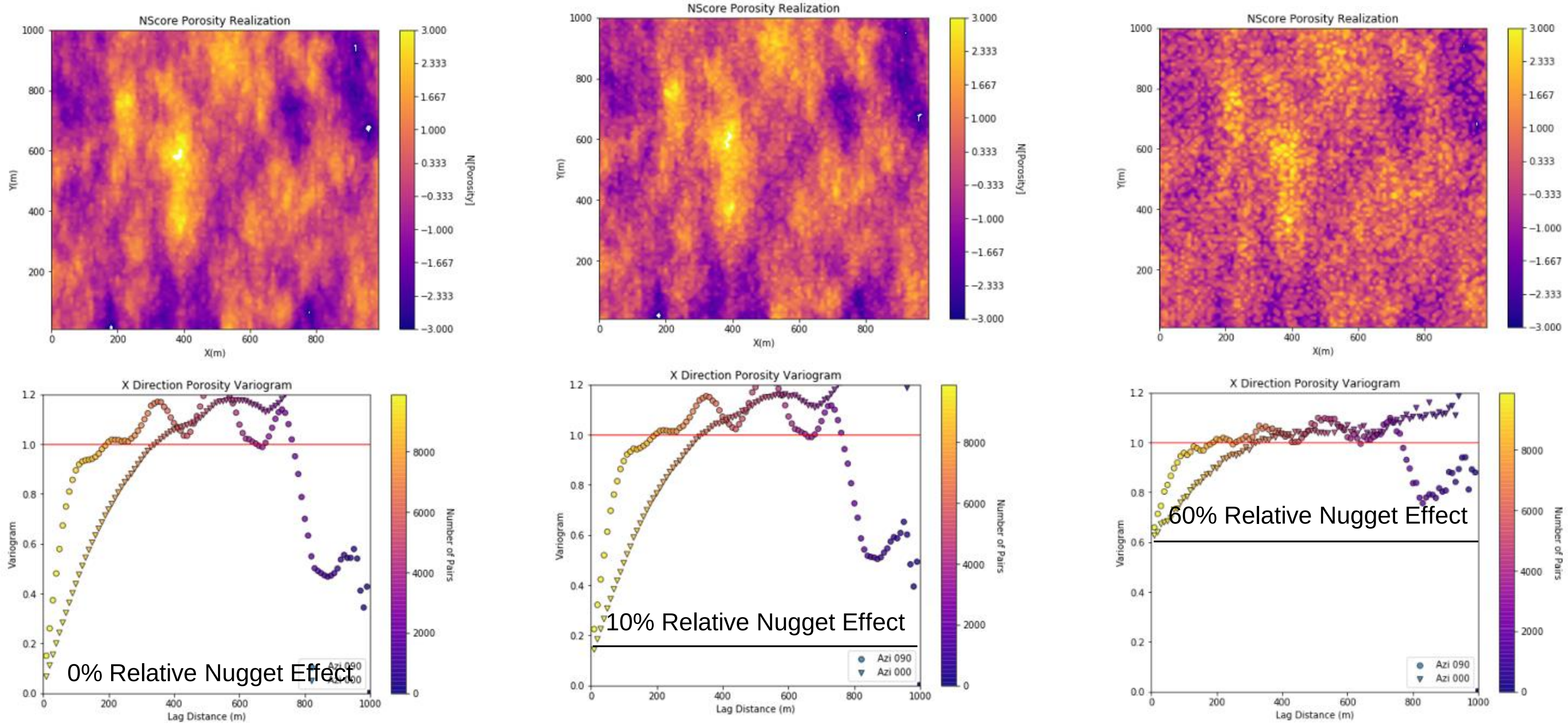
✓Any error in the measurement value or the location assigned to the measurement translates to a higher nugget effect

✓ Sparse data may also lead to a higher than expected nugget effect



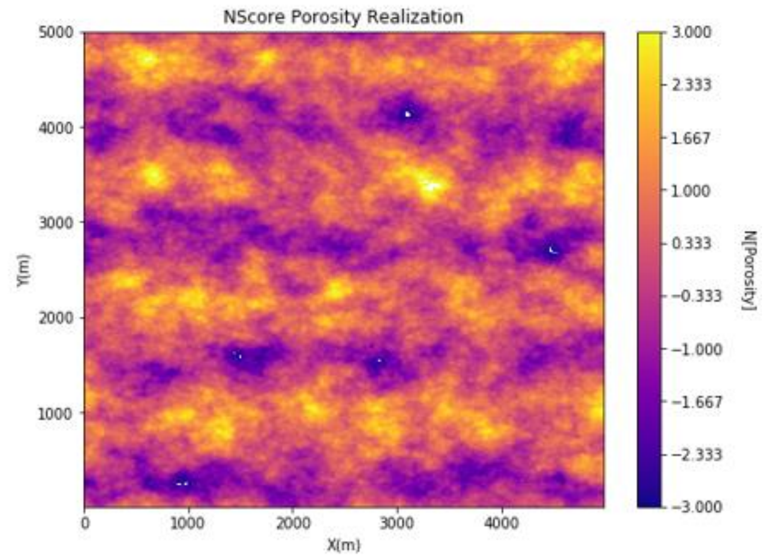
Interpretation Principle Nugget

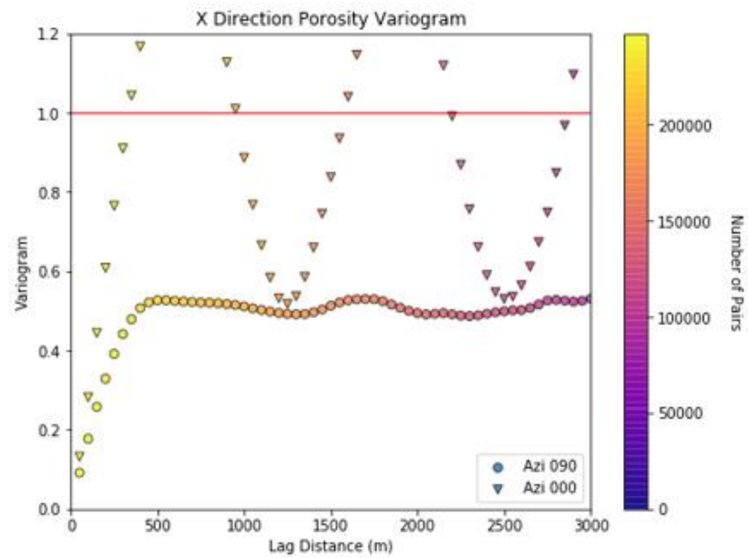
- Variability at a sub-cell scale.
- In this example nugget effect was caused by adding a random value to each cell.



Interpretation Principle Cyclicity

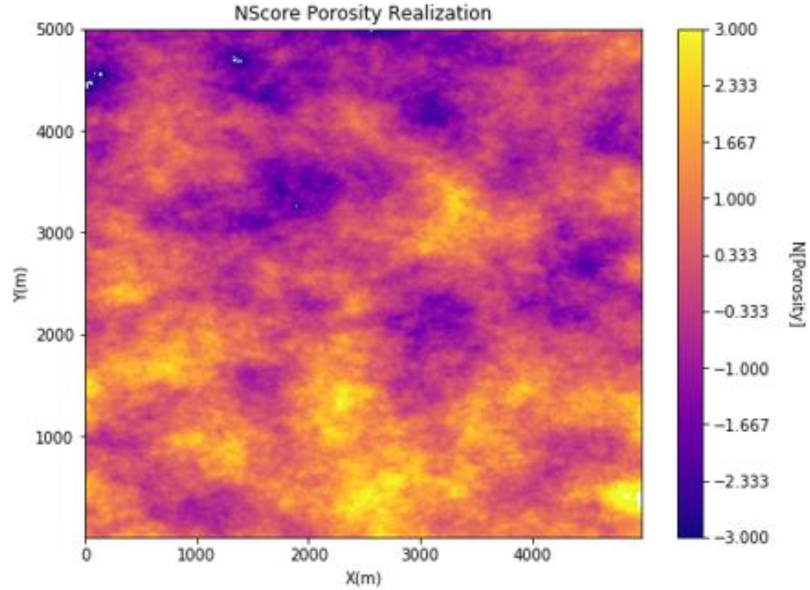
- Cyclicity may be linked to underlying geological directional periodicity
- Could also be noise due to limited data
- Often not modeled with variogram, instead captured with a trend

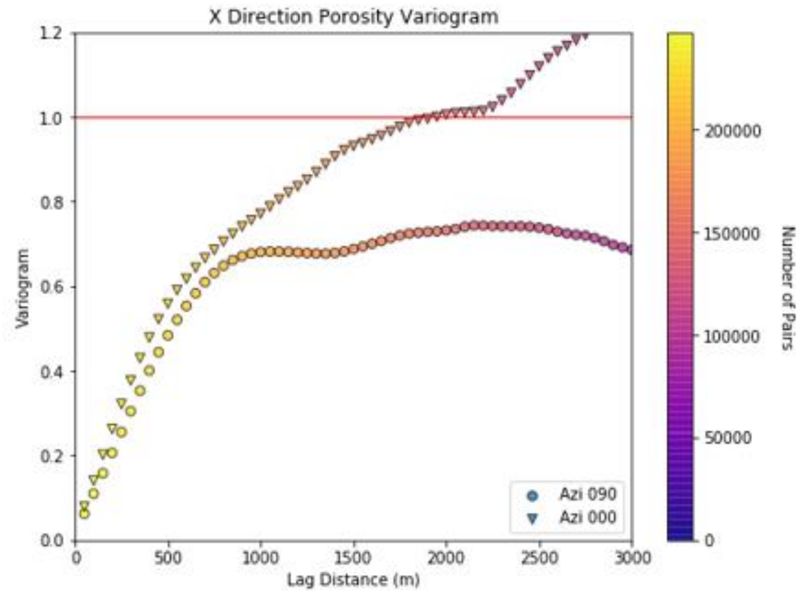




Interpretation Principle Trends

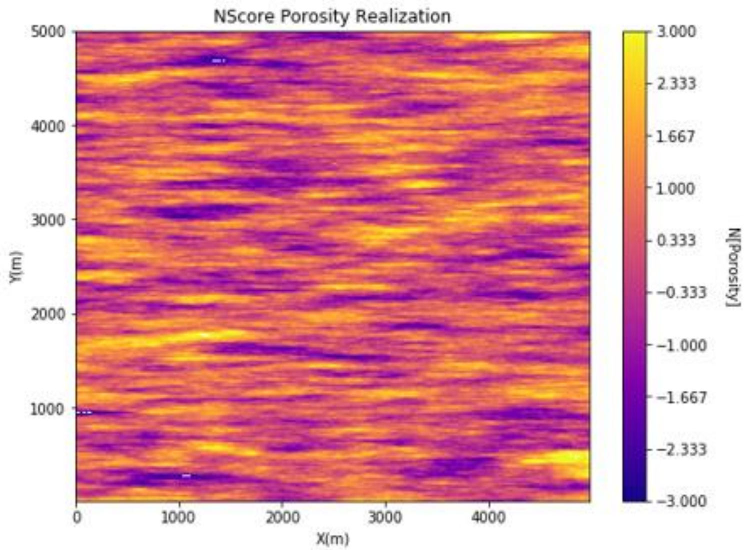
- Variogram rises above the sill
- Indicates a trend (fining upward, fining distal etc.)
- Best to model the trend, remove work with residual and model residual variogram to the sill
- For simulation, theory requires us to model to the sill; therefore, we will not model trend with the variogram

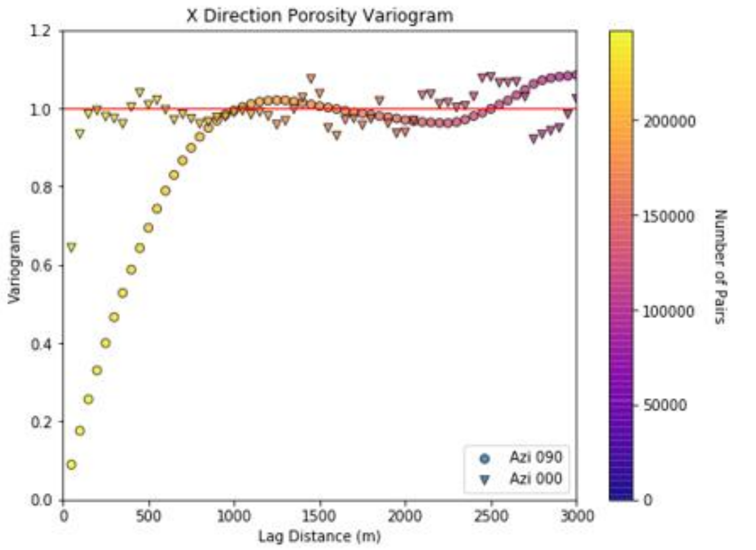




Interpretation Principle
Geometric Anisotropy

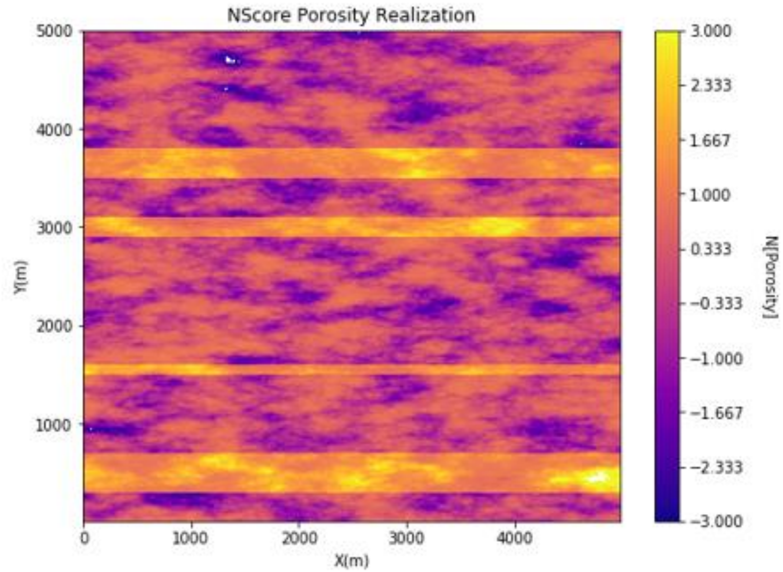
- The range of correlation depends on direction; typically, the vertical range of correlation is much less than horizontal range due to larger lateral distance during deposition
- Geostatistical Walters' Law – same spatial continuity features in all directions with just a different range
- Model the nugget and structure from the vertical direction (or the best informed direction)

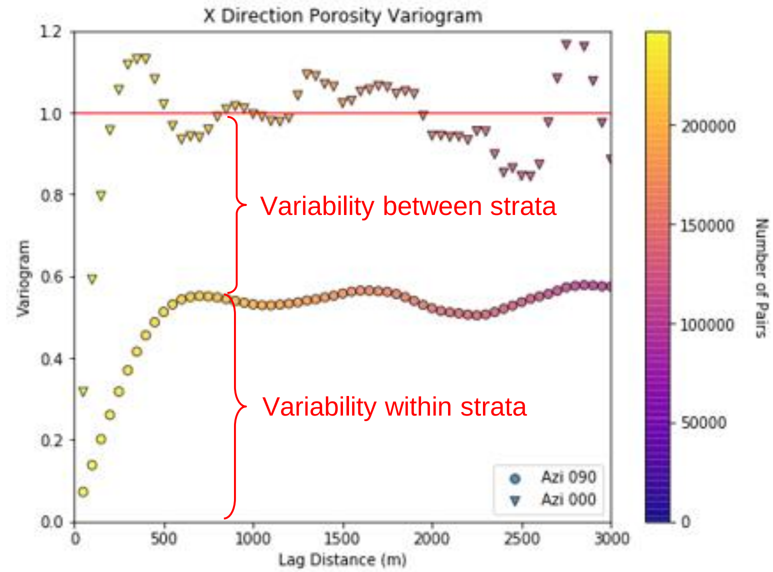




Interpretation Principle
Zonal Anisotropy

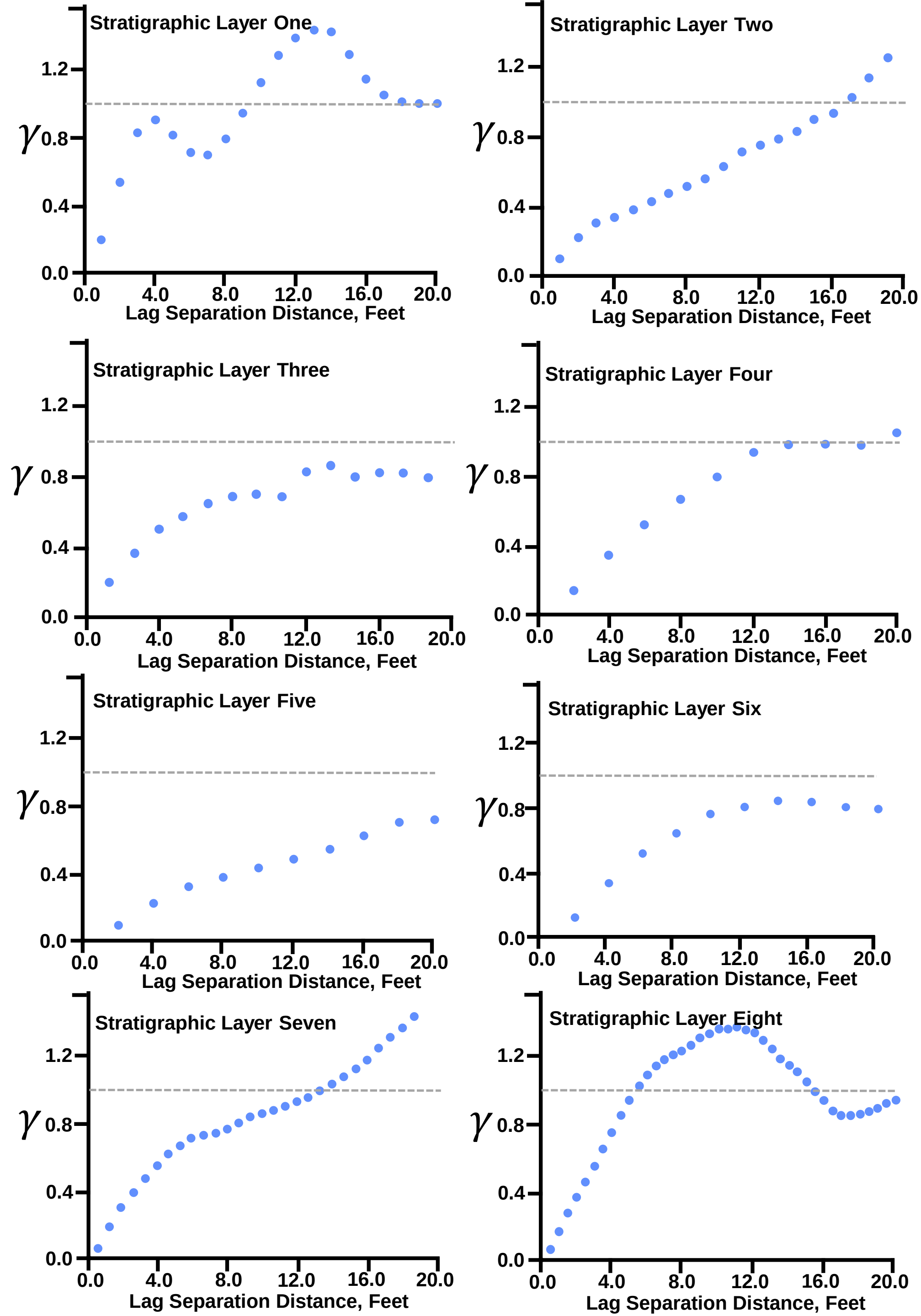
- Variogram does not reach the sill in all directions = apparent sill
- Horizontal variogram doesn't reach the sill:
  - horizontally do not see all the variance within a layer
- Vertical variogram reaches a *lower* sill:
  - likely due to a significant difference in the average value in each well → horizontal variogram has additional between-well variance





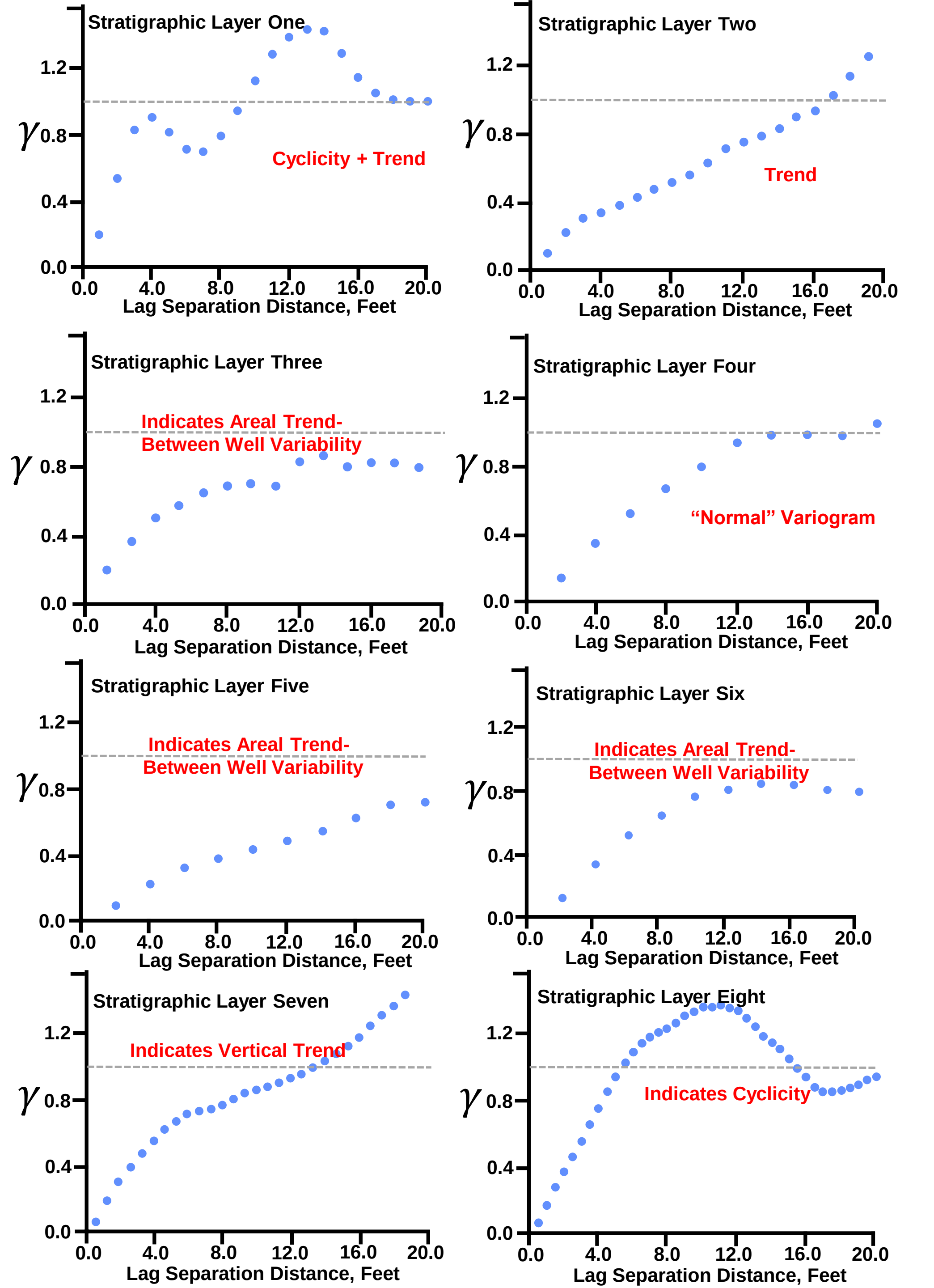
Experimental Variograms Examples

- Superposition of the previous four interpretation principles
  - Trend
  - Cyclicity
  - Zonal anisotropy
  - Geometric anisotropy
- Subsurface heterogeneity is often a combination of features



Experimental Variograms Examples

- Superposition of the previous four interpretation principles
  - Trend
  - Cyclicity
  - Zonal anisotropy
  - Geometric anisotropy
- Subsurface heterogeneity is often a combination of features

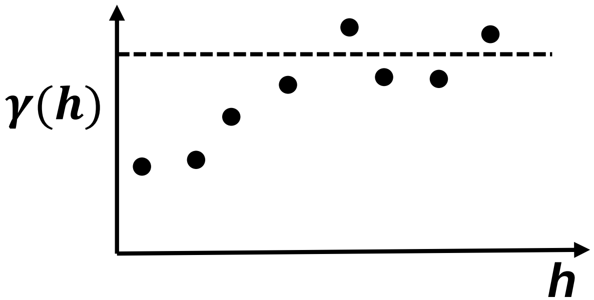


Main Points for Variogram Interpretation

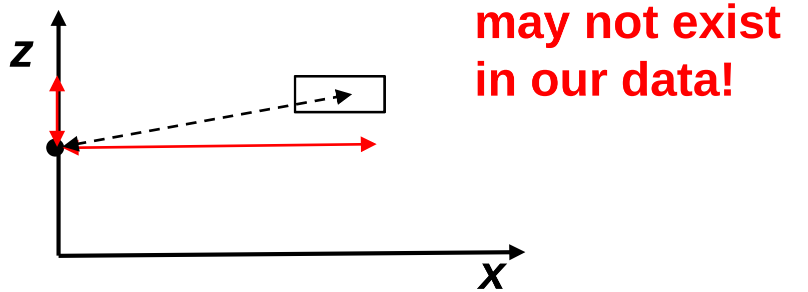
- Variogram is commonly applied to constrain the spatial heterogeneity of continuous reservoir properties
- Variogram can be directly related to aggregation and scaling
- Prior to calculation coordinate and data transformation may be required.
- Interpretation Principles:
  - Trend, Cyclicity, Geometric Anisotropy, Zonal Anisotropy
- Identify and Model Trends Prior to Variogram Calculation
  - residual variogram reaches the sill
- Short scale structure is most important
  - nugget due to measurement error should not be modeled, commonly nugget is not included
  - size of geological modeling cells is a constraint on what can be captured in model!
- Vertical direction may have higher data density, more reliable
  - to establish nugget effect and structure types and then use geometric anisotropy
- Horizontal direction is not well informed
  - take from analog field or outcrop
  - typical horizontal vertical anisotropy ratios

Reasons for Variogram Modeling

1. Need the variogram for *all* distances and directions and data cannot generally provide this – move beyond data..



In between points

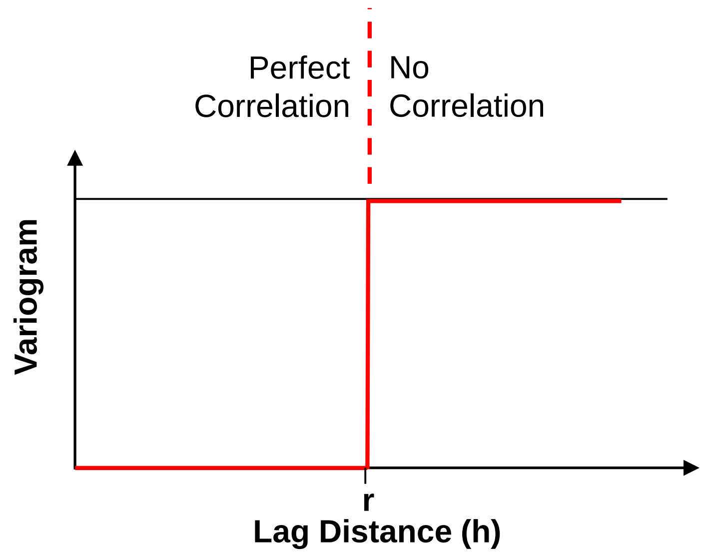


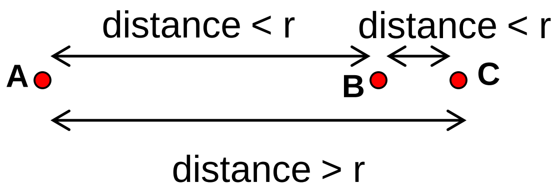
Off diagonal directions

2. Integrate geological knowledge (analog fields, stratigraphic concepts)
3. The variogram model must be positive definite (a legitimate measure of distance), no spatial paradoxes

Reasons for Variogram Modeling

Extreme Example to Demonstrate the Need for Using Positive Definite Variogram





A and B and B and C are perfectly correlated, but A and C are not correlated!

- this is a spatial paradox!

Positive definite variogram models ensure for all possible data and estimate configurations there are no paradoxes.

Reasons for Variogram Modeling

- Cannot just fit a smooth interpolation to the experimental variogram points.

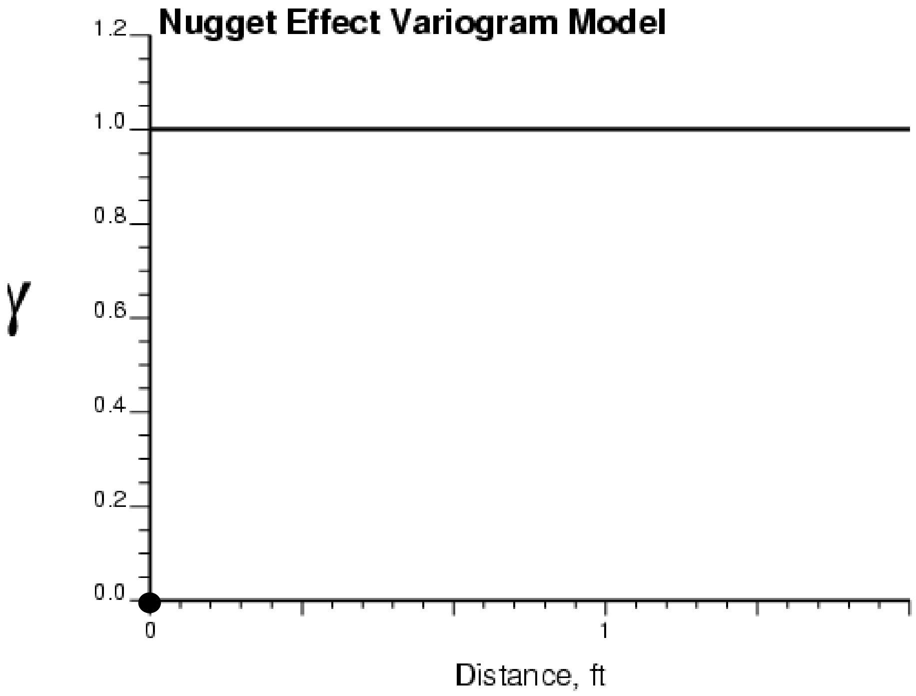
- The variogram will be used in Kriging (next section)

- Kriging uses variogram (covariance) values in a linear system of equations

- Solves for estimate and estimation variance (uncertainty in the estimate)

estimation variance = $\sigma_{x^*}^2 = \sigma_x^2 - \sum_{\alpha=1}^{n} \lambda_\alpha C_x(\mathbf{u}_\alpha - \mathbf{u}_0) \geq 0$

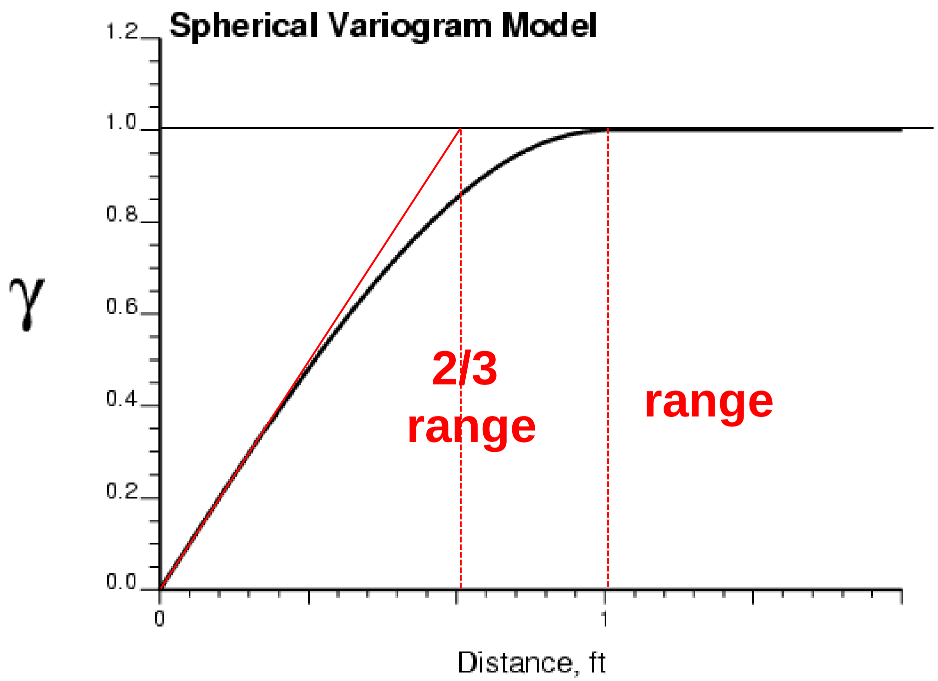
Common Variogram Models



- No spatial correlation
- Should be a small component of the overall variance

$$\gamma(\mathbf{h}) = \begin{cases} 0, & \text{if } \mathbf{h} = \mathbf{0} \\ C(\mathbf{0}), & \text{if } \mathbf{h} > \mathbf{0} \end{cases}$$

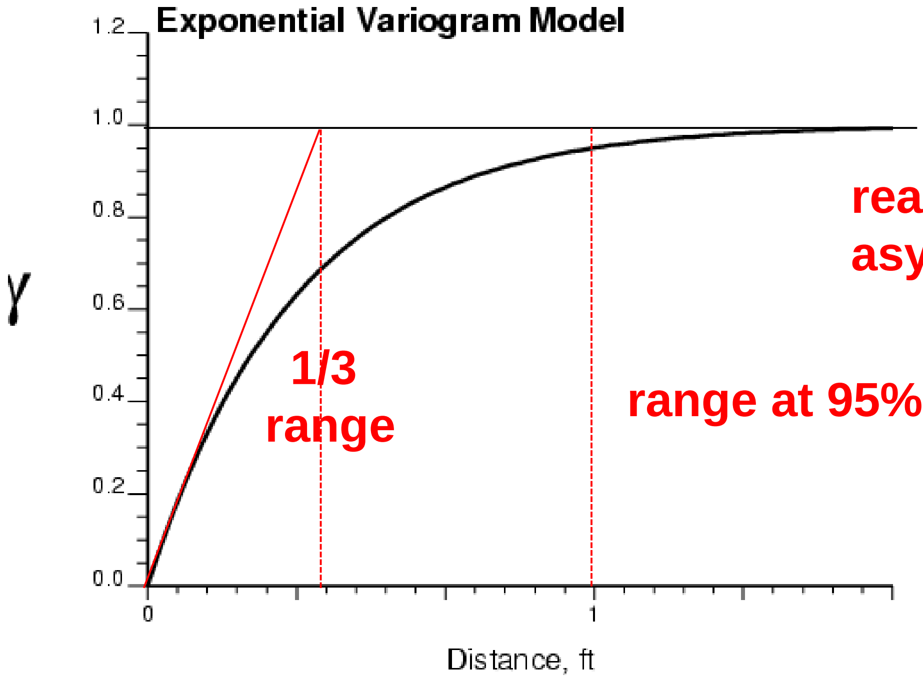
Common Variogram Models



Commonly
encountered
variogram shape

$$\gamma(\mathbf{h}) = C(\mathbf{0}) \cdot Sph\left(\frac{\mathbf{h}}{a}\right) \begin{array}{ll} = c \cdot \left[1.5\left(\frac{\mathbf{h}}{a}\right) - 0.5\left(\frac{\mathbf{h}}{a}\right)^3\right], & \text{if } \mathbf{h} < a \\ = c & \text{if } \mathbf{h} \geq a \end{array}$$

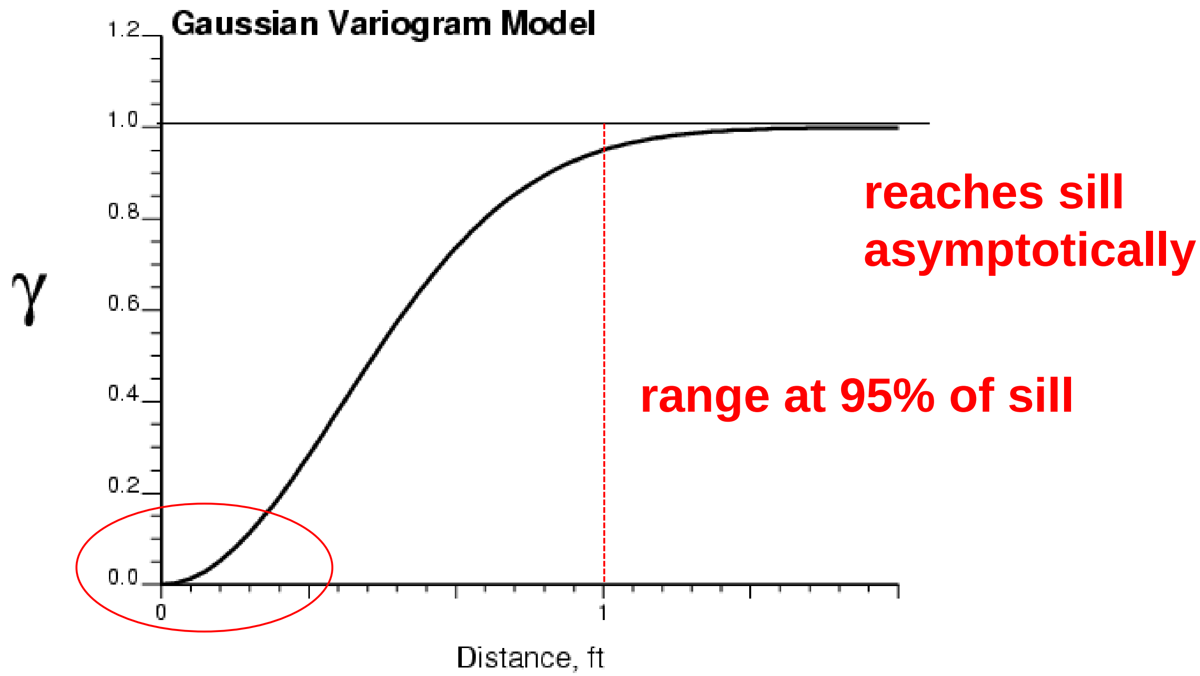
Common Variogram Models



Similar to spherical but rises more steeply and reaches the sill asymptotically

$$\gamma(\mathbf{h}) = C(\mathbf{0}) \cdot Exp\left(\frac{\mathbf{h}}{a}\right) = c \cdot \left[1 - exp\left(-\frac{\mathbf{h}}{a}\right)\right]$$

Common Variogram Models

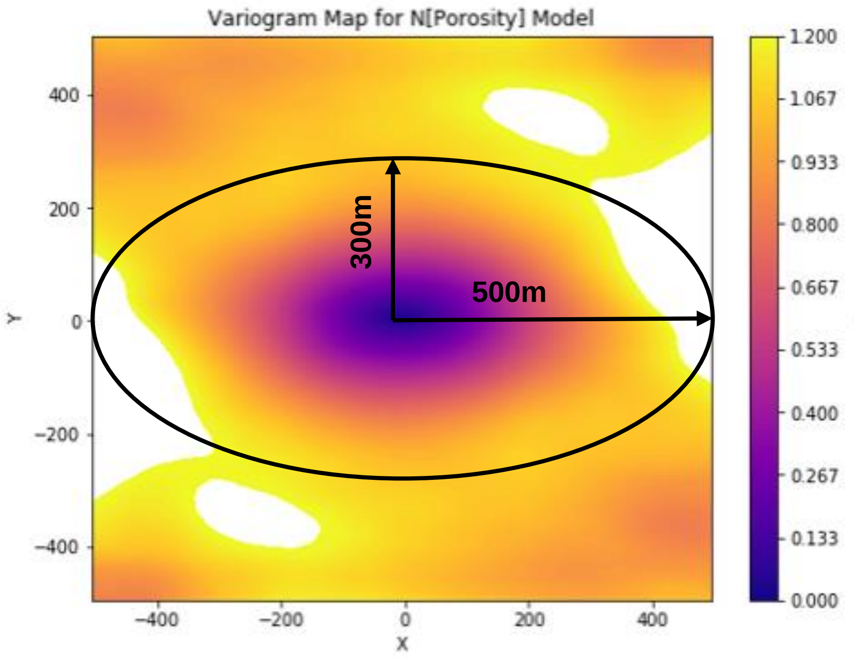


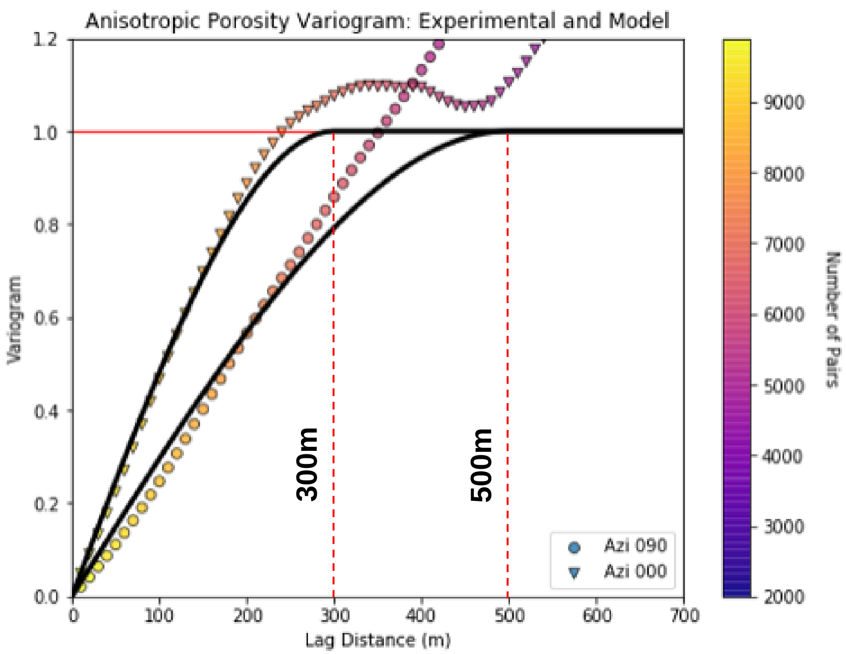
Implies short scale continuity; parabolic behavior at the origin, instead of linear

$$\gamma(\mathbf{h}) = C(\mathbf{0}) \cdot Gauss\left(\frac{\mathbf{h}}{a}\right) = c \cdot \left[1 - exp\left(-\frac{\mathbf{h}^2}{a^2}\right)\right]$$

2D Variogram Models

- Calculate the variogram for all possible distances and directions (variogram map).

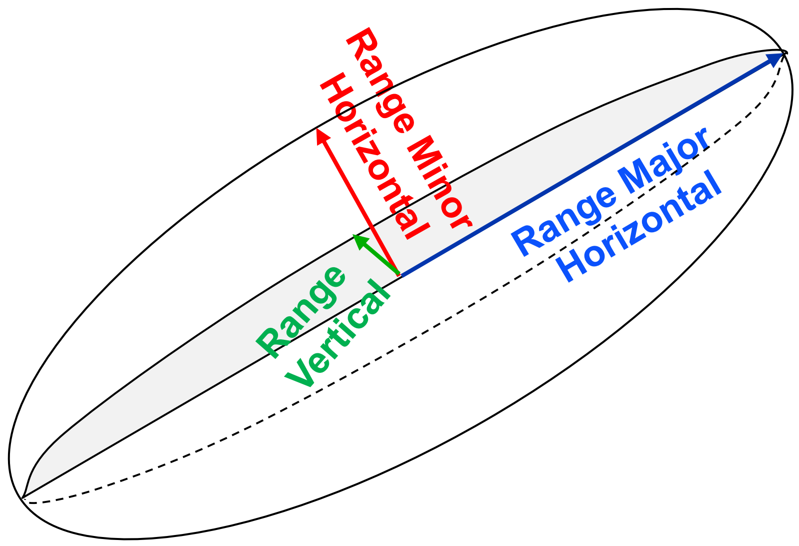




- There is an ellipsoidal variation in continuity (geometric anisotropy):
  - Parameters for a 2D variogram model:
    - direction, major and minor range, type of variogram

2D / 3D Variogram Models

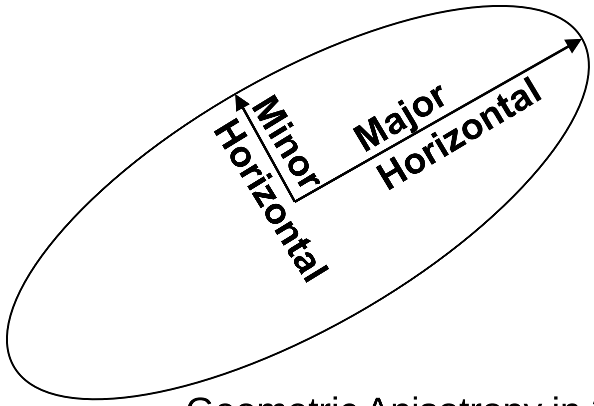
- The variation of range along different directions is modeled using an ellipse in 2D and an ellipsoid in 3D



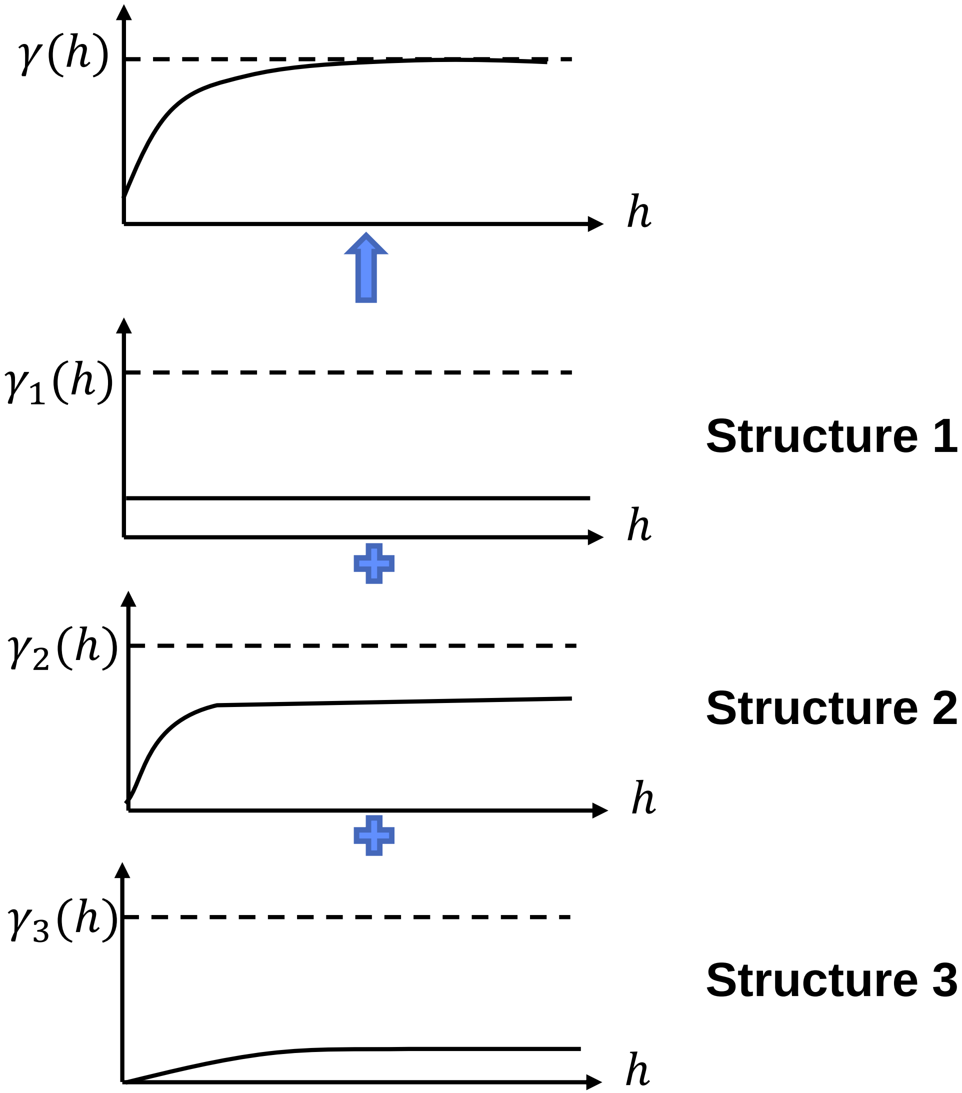
- There is an ellipsoidal variation in continuity (geometric anisotropy):
  - Parameters for a 2D variogram model:
    - direction, dip, major, minor and vertical range, type of variogram

Nested Structures

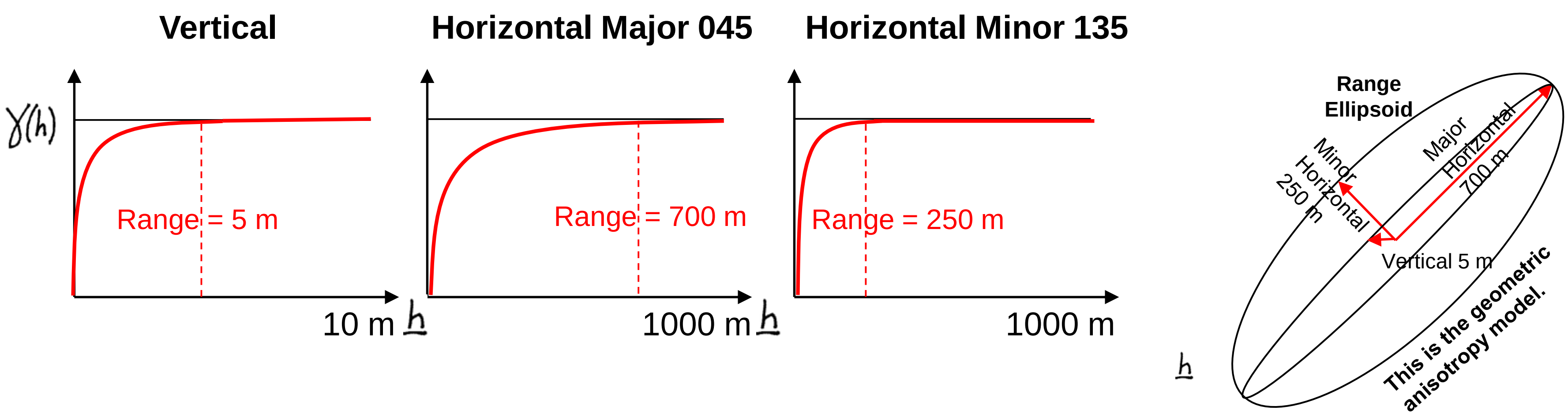
- Addition of positive definite variogram structures is positive definite.
- Each structure covers a proportion of the sill.
- For each structure we can:
  - change the orientation and range in major and minor (geometric anisotropy model).



Geometric Anisotropy in 2D

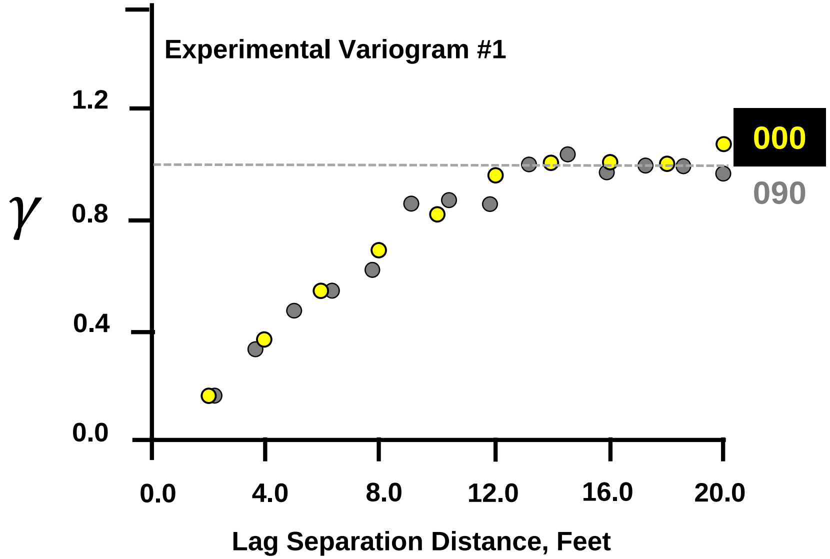


Directional Variograms

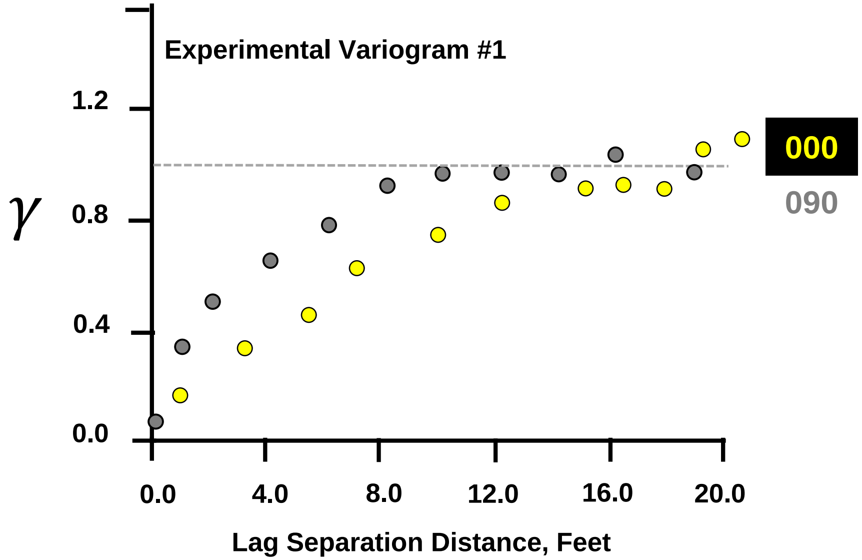


- The variogram models in the principal directions (major horizontal, minor horizontal, and vertical) must be consistent, i.e., same nugget effect and same number and type of structure.
  - this is required so that we can compute variogram values in off-diagonal directions
    - » geometric model, based on an ellipsoid for interpolation between primary directions
  - the responsibility for this is on the person modeling the variogram; most software does not help much
- Basic idea is to explain the total variability by a set of nested structures where each nested structure has different range parameters in different directions:

Hands-on Model Variograms

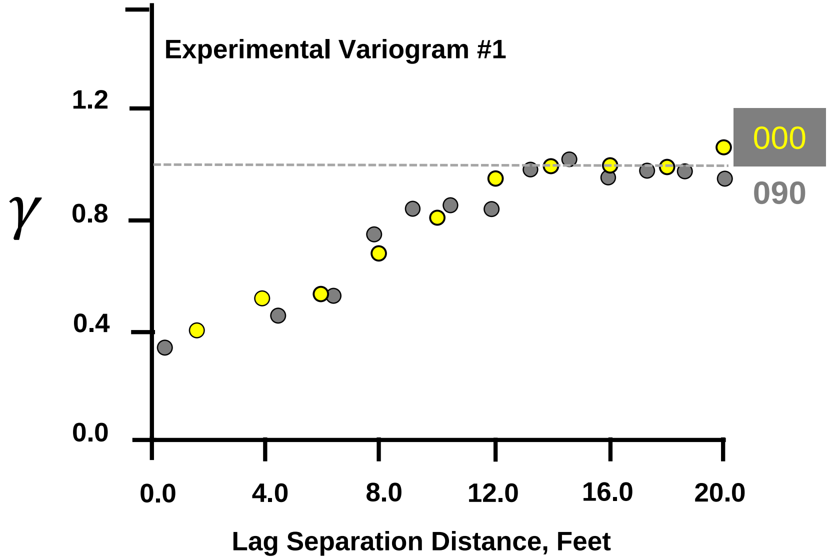


Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1					
2					
3					

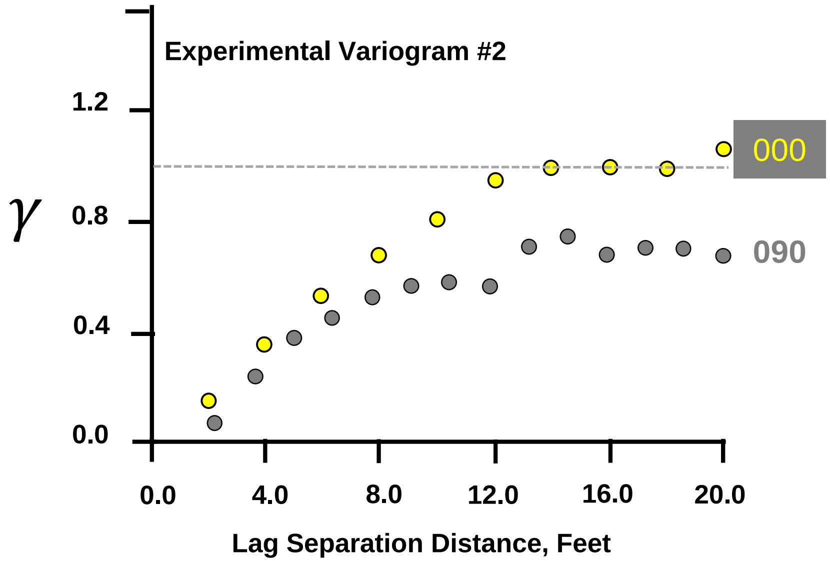


Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1					
2					
3					

Hands-on Model Variograms

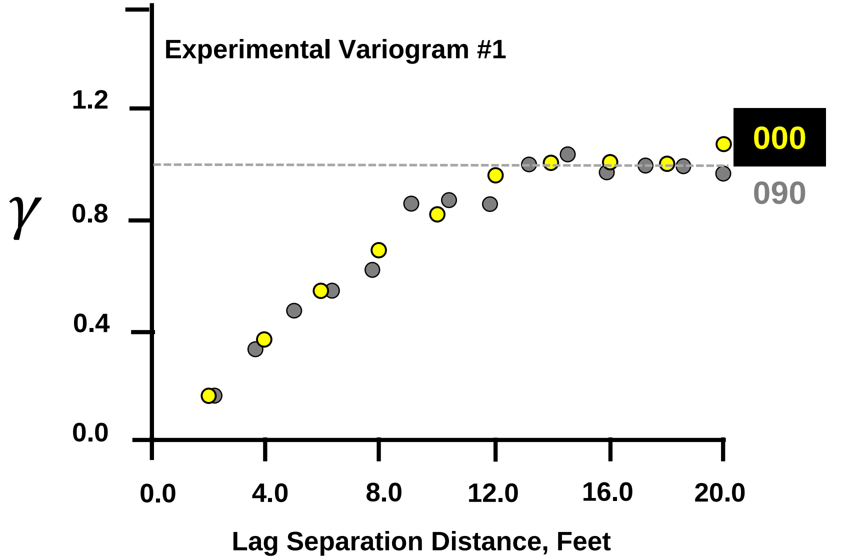


Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1					
2					
3					

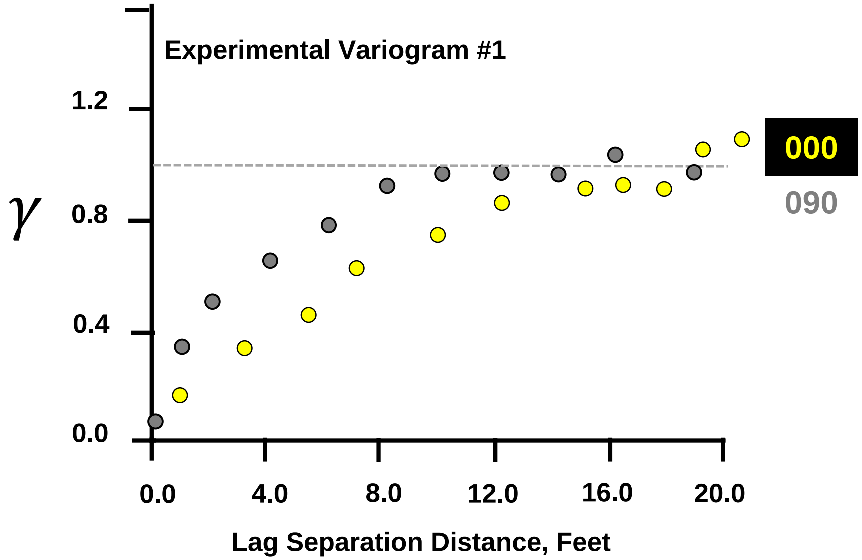


Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1					
2					
3					

Hands-on Model Variograms

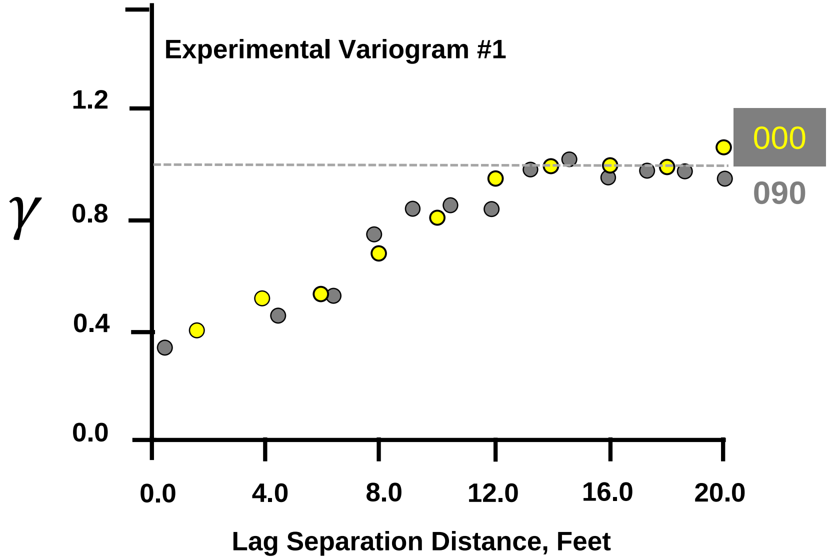


Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1	1.0	Spherical	000	15	15
2					
3					

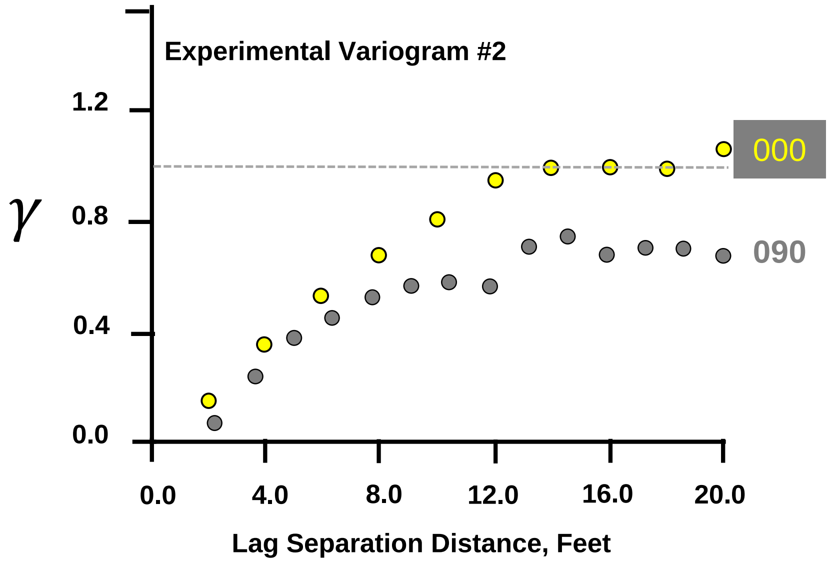


Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1	1.0	Spherical	090	18	10
2					
3					

Hands-on Model Variograms



Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1	0.4	Nugget			
2	0.6	Spherical	000	14	14
3					



Structure	Contribution	Type	Azimuth	Range Major	Range Minor
1	0.7	Spherical	090	14	14
2	0.3	Spherical	090	9999	14
3					

Variogram Modeling

- The following procedure is used to ensure a legitimate model:
  - Pick a single (lowest) isotropic nugget effect from the directionals
  - Choose the same number of variogram structures for all directions based on most complicated directional
  - Ensure that the same contribution of sill parameter is used for all variogram structures in all directions
  - Freely adjust for different range parameter in each direction (geometric anisotropy)
  - Model a zonal anisotropy by setting a very large range parameter in one or more of the principal directions

- Software may assist with auto-fitting, but take ownership, check and integration geological information

Inference in Presence of Sparse Data

- Most often there are inadequate data to infer a reliable horizontal variogram.

- Horizontal wells have not significantly helped with horizontal variogram inference (yet!):
  - *hard* core data and sophisticated well log measurements are rarely collected from horizontal wells
  - horizontal wells rarely track the stratigraphic "time lines"; they typically intersect the formation at some angle that undulates along the length of the wellbore

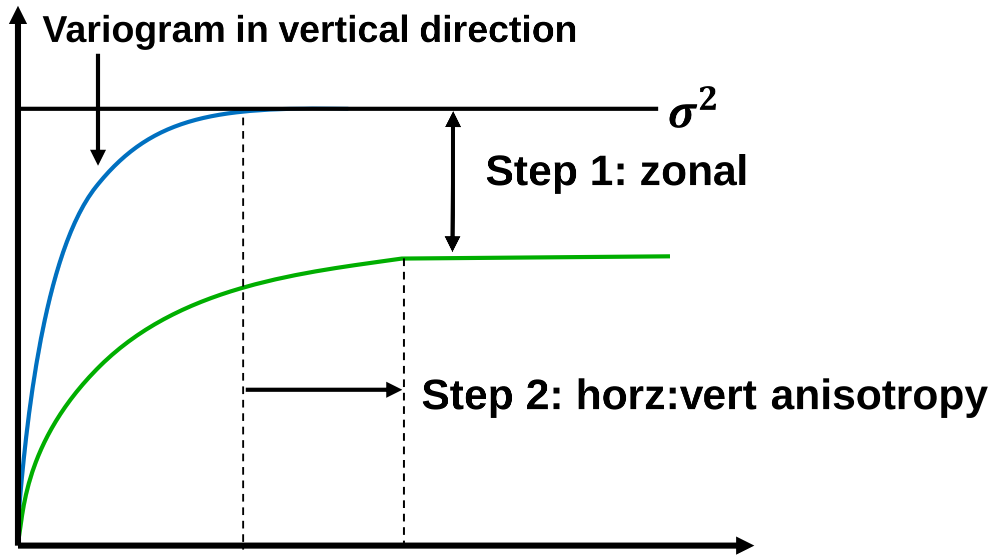
  Horizontal well data will become increasingly important

- At present, we depend on analogue data deemed relevant to the site being considered such as:
  - other, more extensively sampled, reservoirs,
  - geological process simulation, or
  - outcrop measurements

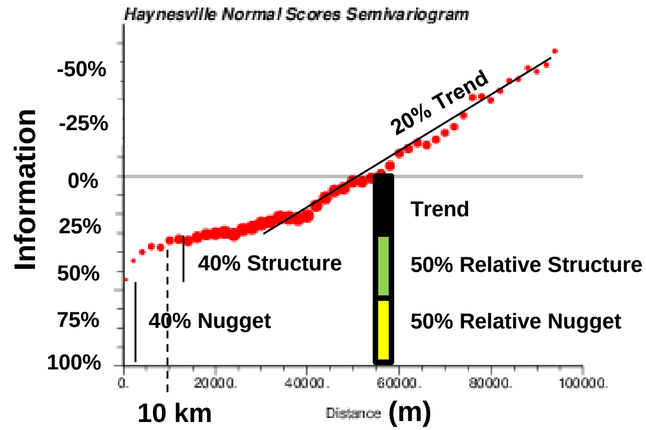
- In all cases, care must be taken to integrate global information from analogues with sparse local data

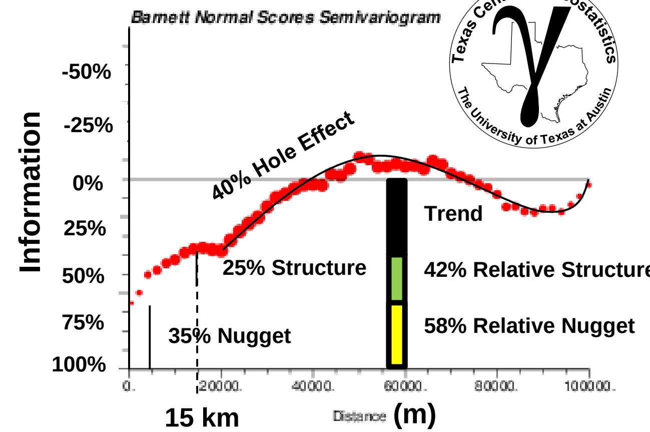
Inference in Presence of Sparse Data

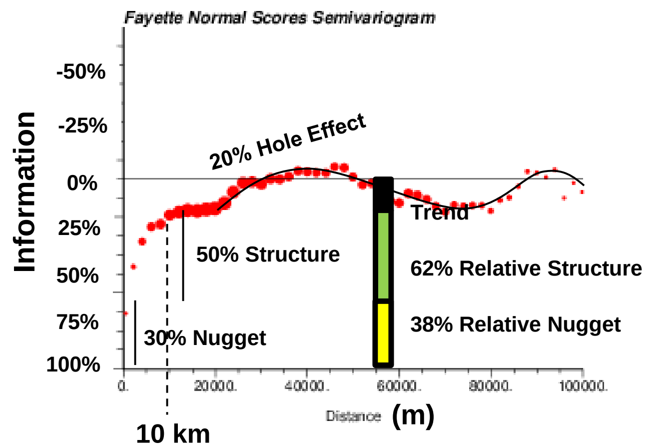
- Calculate experimental variograms: the vertical variogram and the horizontal variogram lags that can be calculated
- Establish zonal anisotropy: estimate the fraction of variance explained “within zones” and “between zones”
- Establish presence of vertical / horizontal trends.

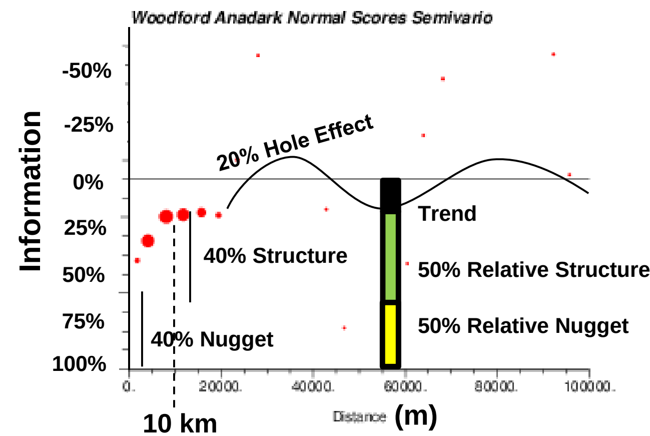


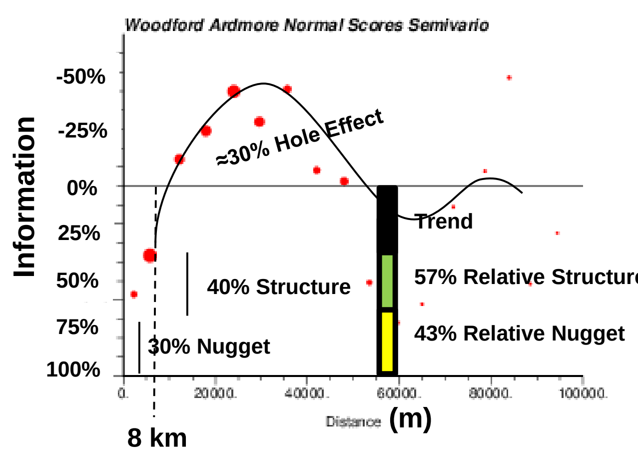
More Examples

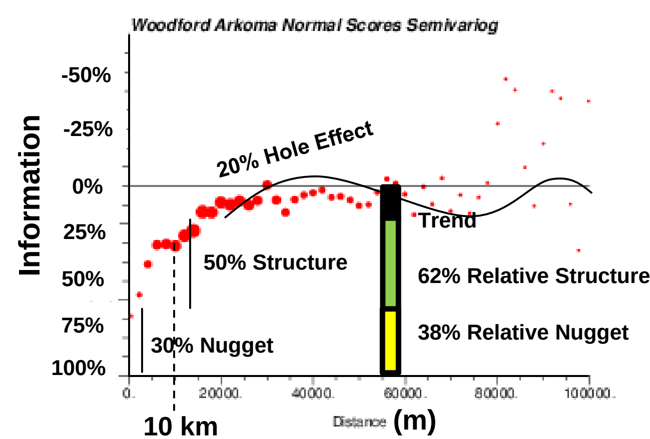












Variogram Modeling Hands-on

The decision of a variogram model impacts the uncertainty in our predictions from a known location.

In the case of a single well the estimation variance (more on this next unit) is equal to the variogram model!

So we can build an Excel spreadsheet to explore the impact of variogram modeling decisions (and trend model) on predictions with uncertainty away from the known location.

Variogram and Trend-based Uncertainty Away from a Single Well

File is Excel/Uncertainty_Away_From_Well_Demo.xls

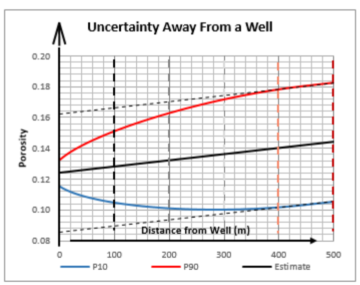
Michael Pyrcz, Geostatistics at Petroleum and Geosystems Engineering, University of Texas at Austin (mpyrcz@austin.utexas.edu)

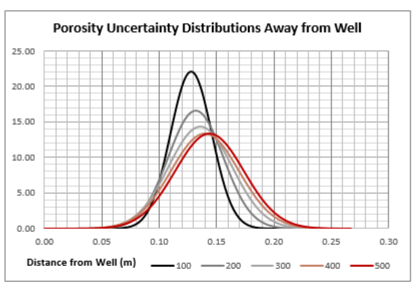
Instructions: set the (1) well porosity value, (2) global porosity variance, (3) trend slope away from the well, and (4) variogram parameterized by the relative nugget effect and spherical range.

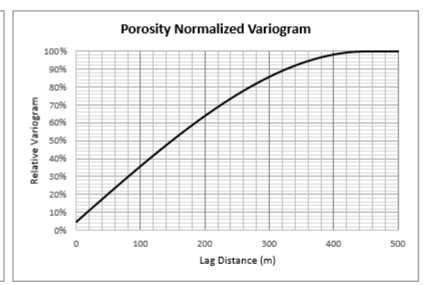
Spatial Model	
Well Value	0.124
Global Var.	0.0009
Trend m	0.00004
Nugget	0.05
Spherical	0.95
Range	450

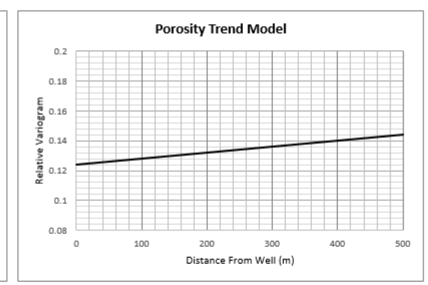
Distance	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130
Estimate	0.124	0.1242	0.1244	0.1246	0.1248	0.125	0.1252	0.1254	0.1256	0.1258	0.126	0.1262	0.1264	0.1266	0.1268	0.127	0.1272	0.1274	0.1276	0.1278	0.128	0.1282	0.1284	0.1286	0.1288	0.129	0.1292
Rel. Var.	5%	7%	8%	10%	11%	13%	14%	16%	18%	19%	21%	22%	24%	25%	27%	29%	30%	32%	33%	35%	36%	38%	39%	41%	42%	44%	45%
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	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	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	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	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	0.17	0.17

Distance	135	140	145	150	155	160	165	170	175	180	185	190	195	200	205	210	215	220	225	230	235	240	245	250	255	260	265
Estimate	0.1294	0.1296	0.1298	0.13	0.1302	0.1304	0.1306	0.1308	0.131	0.1312	0.1314	0.1316	0.1318	0.132	0.1322	0.1324	0.1326	0.1328	0.133	0.1332	0.1334	0.1336	0.1338	0.134	0.1342	0.1344	0.1346
Rel. Var.	46%	48%	49%	51%	52%	54%	55%	56%	58%	59%	60%	62%	63%	64%	65%	67%	68%	69%	70%	71%	73%	74%	75%	76%	77%	78%	79%
St. Dev.	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
P10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
P90	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	0.16	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
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	0.10	0.10	0.10	0.10	0.10	0.10
GlobalP90	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17









Observe the uncertainty model represented by P10, mean and P90 away from the well and the PDFs at distance = 100, 200,..., 500. Modify the variogram model and observe the change. Dashed lines are the maximum uncertianty with no data within the range.

Located at: https://git.io/fNO2I

Distance	Est	St.Dev	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.06	0.06	0.06
100	0.13	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02
200	0.13	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.02	0.03	0.05	0.07	0.10	0.14	0.19
300	0.14	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.03	0.04	0.05	0.07	0.09	0.12	0.16	0.21	0.27	0.34
400	0.14	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.11	0.14	0.18	0.23	0.29	0.36
500	0.14	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.06	0.08	0.10	0.13	0.16	0.21	0.26

Distance	0.06	0.07	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.09	0.09	0.09	0.09	0.10	0.10	0.10	0.10	0.11	0.11	0.11	0.11	0.12	0.12	0.12	0.12	0.13	0.13
100	0.03	0.05	0.08	0.13	0.19	0.29	0.44	0.64	0.92	1.29	1.78	2.40	3.19	4.15	5.29	6.63	8.14	9.81	11.59	13.44	15.29	17.06	18.67	20.05	21.11	21.81	22.11
200	0.25	0.34	0.45	0.60	0.77	1.00	1.27	1.60	1.99	2.45	2.99	3.60	4.30	5.07	5.92	6.84	7.81	8.83	9.87	10.92	11.94	12.93	13.84	14.66	15.35	15.91	16.31
300	0.44	0.55	0.69	0.86	1.06	1.29	1.57	1.89	2.25	2.67	3.13	3.65	4.22	4.84	5.50	6.21	6.94	7.71	8.49	9.27	10.04	10.79	11.50	12.16	12.75	13.27	13.69
400	0.45	0.56	0.69	0.84	1.02	1.23	1.47	1.75	2.07	2.43	2.83	3.27	3.75	4.27	4.83	5.43	6.06	6.71	7.38	8.07	8.75	9.42	10.08	10.70	11.28	11.81	12.28
500	0.33	0.41	0.51	0.63	0.78	0.94	1.14	1.37	1.63	1.92	2.26	2.63	3.05	3.50	4.00	4.54	5.11	5.71	6.34	7.00	7.66	8.33	9.00	9.66	10.29	10.88	11.43

Variogram Modeling in R Demo

Also made a Rmarkdown tutorial for the following workflow example in R with a lot more explanation.

- File name: **variogram_demo.html**

Variogram Analysis in R for Engineers and Geoscientists

Michael Pyrcz, Associate Professor, University of Texas at Austin,

Contacts: Twitter/@GeostatsGuy | GitHub/GeostatsGuy | www.michaelpyrcz.com | GoogleScholar | Book

A tutorial/demonstration of variogram analysis workflow based on the gstat package by Pedesma, E. The docs are at https://cran.r-project.org/web/packages/gstat/index.html. I found Pedesma's Meuse tutorial very helpful (https://cran.r-project.org/web/packages/gstat/vignettes/gstat.pdf). Also, appreciation to Pebesma for assistance through answering questions. For this demonstration we use a 200 well 2D porosity dataset (file: 2D_MV_200Wells.csv) that may be found at https://github.com/GeostatsGuy/GeoDataSets. I used this in my Introduction to Geostatistics undergraduate class (PGE337 at UT Austin) as part of a first introduction to R for the engineering undergraduate students. It is assumed that students have no previous R experience; therefore, all steps of the code are explored and described.

Load the required libraries

```
library(gstat)                              # geostatistical methods by Edzer Pebesma
```

```
## Warning: package 'gstat' was built under R version 3.4.3
```

```
library(sp)                                 # spatial points addition to regular data frames
```

```
## Warning: package 'sp' was built under R version 3.4.3
```

```
library(plyr)                               # manipulating data by Hadley Wickham
```

```
## Warning: package 'plyr' was built under R version 3.4.3
```

If you get an error, you may have to first go to "Tools/Install Packages..." to install these packages. Just type in the names one at a time into the package field and install. The package names should autocomplete (helping you make sure you got the right package name), and the install process is automatic, with the possibility of installing other required dependency packages. Previously I had an issue with packages not being found after install that was resolved with a reboot.

Declare functions

I was surprised that there isn't a built in method to transform a dataframe column or data vector to standard normal, Gaussian with a mean of zero, $\overline{x} = 0.0$ and a standard deviation $\sigma = 1.0$. I found this function by Ashton Shortridge (2008) and included it here. Just apply with the raw data as a vector, x, and it returns an object with the normal score values as a member vector, '[my_transform_object]$nscore'.

```
nscore <- function(x) {                    # by Ashton Shortridge, 2008
  # Takes a vector of values x and calculates their normal scores. Returns
  # a list with the scores and an ordered table of original values and
  # scores, which is useful as a back-transform table. See backtr().
  nscore <- qqnorm(x, plot.it = FALSE)$x  # normal score
  trn.table <- data.frame(x=sort(x),nscore=sort(nscore))
  return (list(nscore=nscore, trn.table=trn.table))
}
```

Variogram Modeling in R Demo

```r
# Variogram Analysis in R for Engineers and Geoscientists New to R
# Michael Pyrcz, University of Texas at Austin, Twitter @GeostatsGuy

# This will be used in my Introduction to Geostatistics Undergraduate Class
# It is assumed that students have no previous R experience.
# This utilizes the gstat library by Edzer Pedesma, appreciation to Dr. Pedesma for assistance.

# Load the required libraries, you may have to first go to "Tools/Install Packages..." to install these first
library(gstat)                         # geostatistical methods by Edzer Pebesma
library(sp)                            # spatial points addition to regular data frames
library(plyr)                          # splitting, applying and combining data by Hadley Wickham
library(ggplot2)

# Declare functions, this function completes standard normal transform on a dataset

nscore <- function(x) {                # written by Ashton Shortridge, May/June, 2008
  # Takes a vector of values x and calculates their normal scores. Returns
  # a list with the scores and an ordered table of original values and
  # scores, which is useful as a back-transform table. See backtr().
  nscore <- qqnorm(x, plot.it = FALSE)$x  # normal score
  trn.table <- data.frame(x=sort(x),nscore=sort(nscore))
  return (list(nscore=nscore, trn.table=trn.table))
}

# Set the working directory, I always like to do this so I don't lose files and to simplify subsequent read and writes
setwd("C:/PGE337")

# Read the data table from a comma delimited file - data on GitHub/GeostatsGuy/GeoDataSets
mydata = read.csv("2D_MV_200wells.csv")          # read in comma delimited data file

# Let's visualize the first several rows of our data so we can make sure we successfully loaded it
head(mydata)                           # show the first several rows of a data table in the console

# The columns are variables with variable names at the top and the rows are samples

# Convert the dataframe to a spatial points dataframe
class(mydata)                          # confirms that it is a dataframe
coordinates(mydata) = ~X+Y             # indicate the X, Y spatial coordinates
summary(mydata)                        # confirms that it is now a spatial points dataframe
head(coordinates(mydata))              # check the first several coordinates

# Normal scores transform of the porosity data
npor.trn = nscore(mydata$porosity)     # normal scores transform
mydata[["NPorosity"]]<-npor.trn$nscore         # append the normal scores transform into the spatial data table
head(mydata)
```

In RStudio load the file variogram_demo.R. Let's step through it together.

Variogram Modeling in R Demo

- Workflow steps:

1. Load Data file "2D_MV_200Well.csv"
2. Transform Porosity to Standard Normal
3. Check Univariate:
    1. PDF and CDF Porosity
    2. PDF and CDF N[Porosity]
4. Check Spatial
    1. Bubble plot
    2. Location map
5. Isotropic Variogram
    1. Calculate
    2. Model
6. Variogram Map
7. Anisotropic Variogram
    1. Calculate
    2. Model
8. Automatic Fitting
9. Custom Variogram Plot

Check for outliers, shape

Variogram search parameters

Spatial continuity?

Anisotropy?

Final Model

Improved plot for interpretation

Variogram Modeling in R Demo

Comments to assist with using R for variogram calculation and modeling.

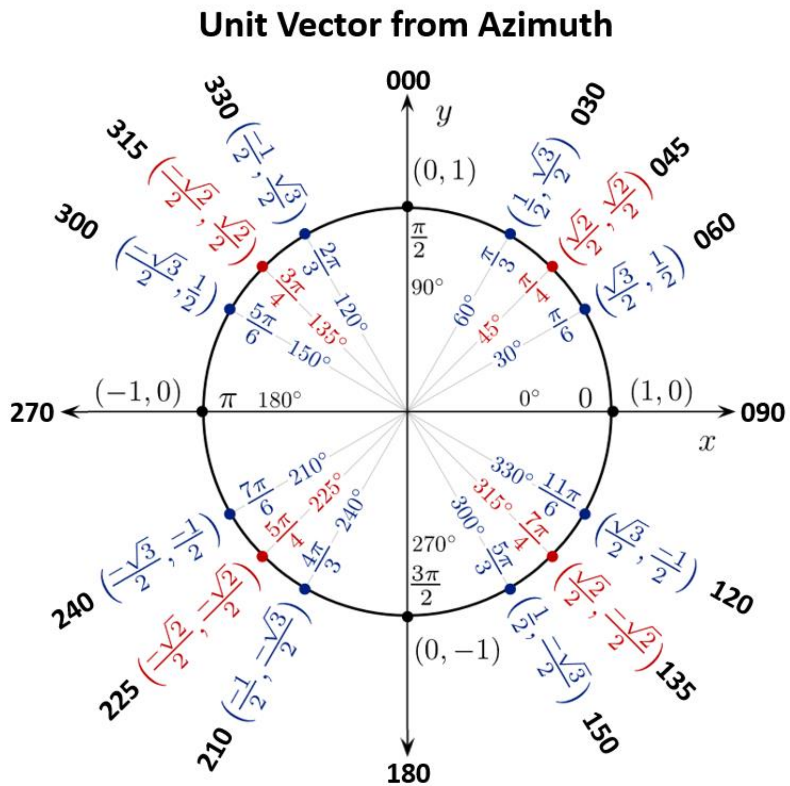
1. Built in variogram plots are inflexible and do not plot the sill, but are easy.

```
>plot(directional_experimental_variogram,variogram_model)
```

will plot your calculated experimental points and the projection of the variogram model in the same direction automatically.

If you use the custom variogram plot method that I wrote out, then you have to get your own projection of the variogram model in the correct direction using the unit vector.

Figure shows unit vector for various azimuths in 2D. Add a extra ‘,0’ to the vector to indicate 3D vector with no vertical component.



Modified from https://upload.wikimedia.org/wikipedia/commons/4/4c/Unit_circle_angles_color.svg
Michael Pyrcz, University of Texas at Austin

```
>unit_vector = c(0,1,0)    # for azimuth 090
>variogramLine(variogram_model,max_lag_distance,min_lag_distance,number_points,unit_vector,covariance=FALSE)
```

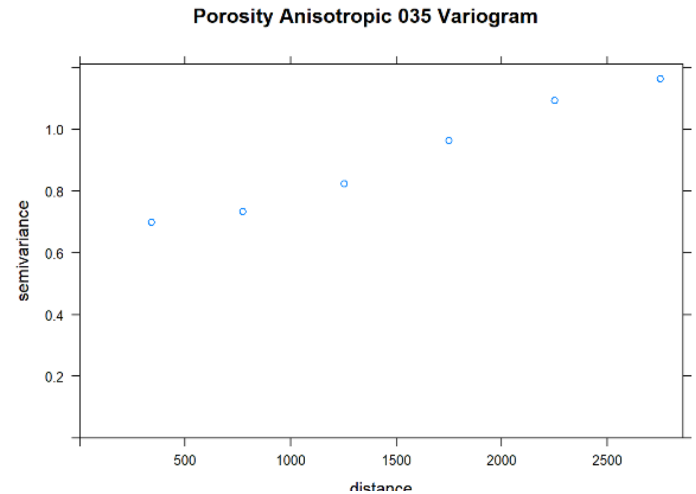
Remember your variogram model is a 2D model that could be evaluated for any direction and distance combination.

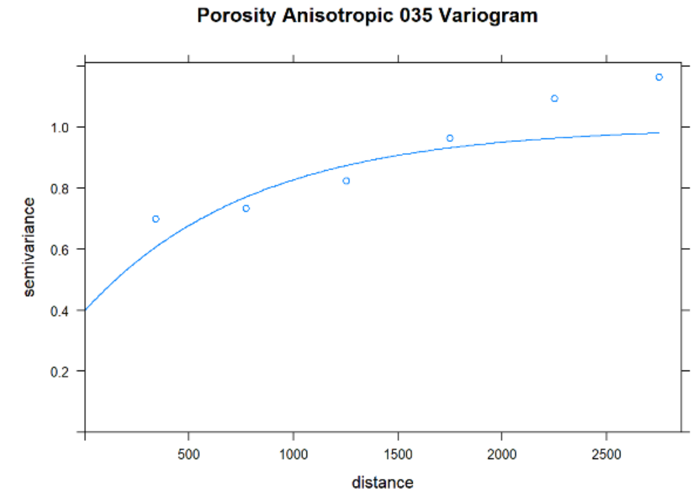
Variogram Modeling in R Demo

Comments to assist with using R for variogram calculation and modeling.

2. Remember this is a modeling exercise.
   - Calculate the variogram in the isotropic and anisotropic cases (direction 090 and 000).
   - Interpret the variogram structures.
   - Fit a nested (nugget and other structure), 2D variogram model.
     - There should be some iteration

- Model to a sill of one, since the variogram is of the N[0,1] transformed porosity variable; therefore, variance = 1.0.
- The model should be plotted and as a line with the experimental points. If the sill is not included, then draw the sill on the plot (if you can't indicate the sill in the caption). The model should be listed as parameters
  - Nugget
  - Contribution of other structure
  - Direction of major horizontal
  - Range in major horizontal
  - Range in minor horizontal
  - Range in R is major, then anisotropy ratio (minor range/major range). E.g. major = 800m, anis1 = 0.5; therefore, minor = 400m.






```
##   model psill range ang1 anis1
## 1   Nug   0.4     0    0   1.0
## 2   Exp   0.6   800   35   0.5
```

Variogram Modeling in R Demo

I gave you some good search parameters to calculate the experimental variograms. Try out the following:

```
por.vg.035 = variogram(NPorosity~1,mydata,cutoff = 3000,width =500,alpha = 35.0,tol.hor=22.5) # azimuth 035

por.vg.125 = variogram(NPorosity~1,mydata,cutoff = 3000,width =500,alpha = 125.0,tol.hor=22.5) # azimuth 125

plot(por.vg.035,main="Porosity Anisotropic 035 Variogram")
```

1. Decrease and increase the lag bins (width parameter in variogram function).

2. Decrease and increase the angle tolerance (tol.hor parameter in the variogram function).

3. Change the azimuths (still 90 degrees difference) (alpha parameter in the variogram function).

Variogram Modeling in Python / GSLIB Hands-on

I gave you some good search parameters to calculate the experimental variograms. Try out the following:

Open Jupyter Notebook form Anaconda3 (64-bit).

Navigate to and load the **Variogram.ipynb** file.

Variogram Calculation in Python for Engineers and Geoscientists

Michael Pyrcz, Associate Professor, University of Texas at Austin

Contacts: Twitter/@GeostatsGuy | GitHub/GeostatsGuy | www.michaelpyrcz.com | GoogleScholar | Book

This is a tutorial for / demonstration of **spatial variogram calculation in Python with simple wrappers and reimplementations of GSLIB: Geostatistical Library methods** (Deutsch and Journel, 1997). Variogram calculation is a valuable method for quantifying spaital continuity. We can interpret the resulting experimental variograms and then infer valid variogram models for use with spatial estimation and simulation.

We will demonstration calculation of variograms on regular and irregular spaced data.

This exercise demonstrates the cell-based declustering approach in Python with wrappers and reimplimentation of GSLIB methods. The steps include:

1. generate a 2D sequential Guassian simulation using a wrapper of GSLIB's sgsim method
2. calculatate the variogram map and anisotropic experimental variograms in the x and y directions with gam
3. visualize the experimental variograms
4. fit a positive definite variogram model with nested know licit variogram structures
5. resample to form a nonuniformaly sampled dataset
6. calculate the variogram map, isotropic and directional experimental variograms with gamv
7. fit a positive definite variogram model with nested know licit variogram structures

To accomplish this I have provide wrappers or reimplementation in Python for the following GSLIB methods:

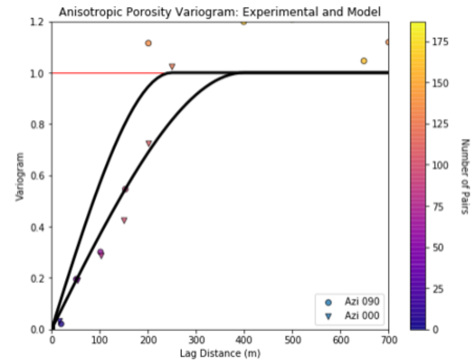
1. sgsim - sequantial Gaussian simulation limited to 2D and unconditional
2. hist - histograms plots reimplemented with GSLIB parameters using python methods
3. locmap - location maps reimplemented with GSLIB parameters using python methods
4. pixelplt - pixel plots reimplemented with GSLIB parameters using python methods
5. locpix - my modification of GSLIB to superimpose a location map on a pixel plot reimplemented with GSLIB parameters using Python methods
6. affine - affine correction adjust the mean and standard deviation of a feature reimplemented with GSLIB parameters using Python methods
7. varmap - vairogram map
8. gam -regularly sampled data variograms
9. gamv - irregularly sampled data variograms
10. nscore - normal score transform (data transformation to Gaussian with a mean of zero and a standard deviation of one)

These methods are all in the functions declared upfront. To run this demo all one has to do is download and place in your working directory the following executables from the GSLIB/bin directory:

Let's step through each code block together.

and visualize the resulting anisotropic variogram model with the experiemental variogram points.

In [22]:
```
lags = np.array([lag_090,lag_000,lag_090_mod,lag_000_mod])
gammas = np.array([gamma_090,gamma_000,gamma_090_mod,gamma_000_mod])
npairs = np.array([npair_090,npair_000])
vtypes = [0,0,1,1]
names = ['Azi 090','Azi 000','Model','Model']
vargplts(lags,gammas,npairs,vtypes,names,0,700,0,1.2,1,"Anisotropic Porosity Variogram: Experimental and Model",cmap)
```



Out[22]: `<matplotlib.collections.PathCollection at 0x19eef6f67b8>`

More Exercises

There are so many more exercises and tests that one could attempt to gain experience with spatial continuity modeling in Python. I'll end here for brevity, but I invite you to continue. Consider, on your own apply other data sets or attempting calculating variograms in different directions and fitting with different models. I hope you found this tutorial useful. I'm always happy to discuss geostatistics, statistical modeling, uncertainty modeling and machine learning,

Michael

Variogram Modeling in Python / GSLIB Hands-on

I gave you some good search parameters to calculate the experimental variograms. Try out the following:

In [19]:
```
nlag = 70; lag = 50; azi = 90; atol = 45; standardize = 0        # set the variogram parameters
# calculate the anisotropic variogram 090
lag_000,gamma_000,npair_000 = gamv_2d(rand_sample,'X','Y','Porosity',nlag,lag,azi,atol,standardize)
nlag = 70; lag = 50; azi = 0; atol = 45; standardize = 0        # set the variogram parameters
# calculate the anisotropic variogram 000
lag_090,gamma_090,npair_090 = gamv_2d(rand_sample,'X','Y','Porosity',nlag,lag,azi,atol,standardize)
```

1. Decrease and increase the lag bins (step parameter in code block 19).

2. Decrease and increase the angle tolerance (atol parameter in code block 19).

3. Change the azimuths (still 90 degrees difference) (azi parameter in the variogram function).

Review of Main Points

- Variogram is very important in geostatistical study; Measure of geological distance
- Initial coordinate and data transformation may be required.
- Interpretation Principles:
  - Trend
  - Cyclicity
  - Geometric Anisotropy
  - Zonal Anisotropy
- Short scale structure is most important
  - nugget due to measurement error should not be modeled
  - size of geological modeling cells
- Vertical direction is typically well informed
  - can have artifacts due to spacing of core data
  - handle vertical trends and areal variations
- Horizontal direction is not well informed
  - take from analog field or outcrop
  - typical horizontal vertical anisotropy ratios

Variogram Modeling
New Tools

Topic	Application to Subsurface Modeling
Model and Remove Trend	Remove a trend from the data and model the variogram of the residual. *Residual variogram provides quantification of trend removal, experimental variogram should each and not exceed the sill.*
Variogram Interpretation	Calculate and interpret the variograms from available data sets. *Interpret superimposed features such as trend, geometric anisotropy, zonal anisotropy, and cyclicity.*
Variogram Modeling	Model the experimental variogram with positive definite variogram structures. *Provides a model of degree of correlation and distance to reach no spatial correlation away from available spatial data.*

Probability and Statistics
What should you learn from this lecture?

Spatial Statistics

- **Variogram Interpretation**
- **Variogram Modeling**

Prerequisites

Introduction

Probability Theory

Representative Sampling

Spatial Data Analysis

Spatial Estimation

Stochastic Simulation

Uncertainty Management

Machine Learning

Additional Information:

Gringarten, E, and Deutsch, C.V., 2001, Teacher's Aide Variogram Interpretation and Modeling, Mathematical Geology, V 35, No. 4.