ArcGIS for Geostatistical Analyst: An Introduction

Steve Lynch and Eric Krause

Redlands, CA.

Outline

- What is geostatistics?
- What is Geostatistical Analyst?
- Spatial autocorrelation
- Geostatistical Wizard and geoprocessing tools
- Where is it used?
- Demonstrations
- Tips and Tricks
- Conclusion
- Questions

Sessions of note…

Tuesday

- Kriging: An Introduction to Concepts and Applications (10:00-11:00 SDCC Rm02)
- **ArcGIS for Geostatistical Analyst: An Introduction (1:00-2:00 SDCC Rm30A)**
- Examples of Geostatistics in Practice (1:00-2:00 SDCC Rm 31A)
- Empirical Bayesian Kriging and EBK Regression Prediction in ArcGIS (2:30-3:15 SDCC Demo Theater 12)
- **ArcGIS for Geostatistical Analyst: An Introduction (4:00-5:00 SDCC Rm30A)**

Wednesday

- Interpolating Surfaces in ArcGIS (10:00-11:00 SDCC Rm30C)
- Surface Interpolation in ArcGIS (11:15-12:00 SDCC Demo Theater 10)
- Examples of Geostatistics in Practice (2:30-3:30 SDCC Rm 29D)
- Kriging: An Introduction to Concepts and Applications (4:00-5:00 SDCC Rm08)

Thursday

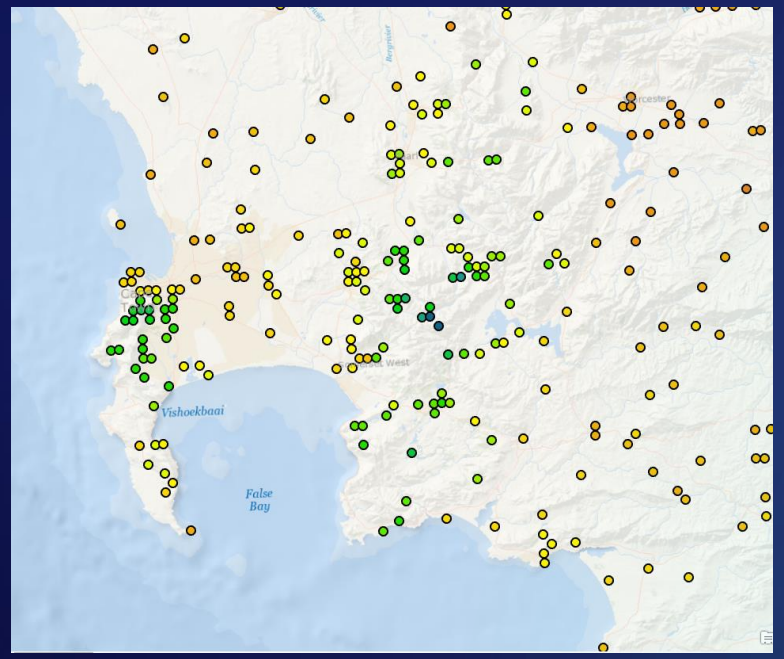
- Polygon-to-Polygon Predictions Using Areal Interpolation (11:15-12:00 SDCC Demo Theater 10)
- Interpolating Surfaces in ArcGIS (4:00-5:00 SDCC Rm31A)

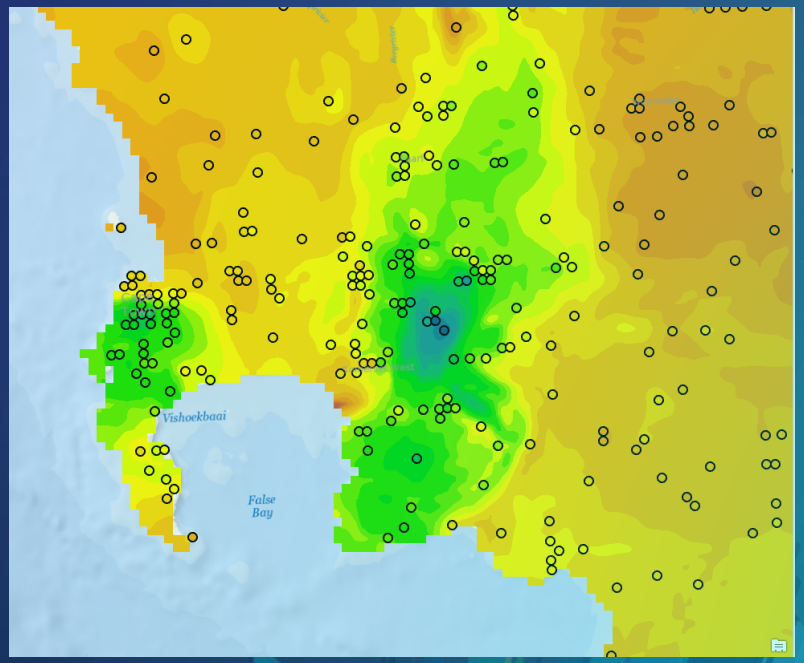
What is geostatistics?

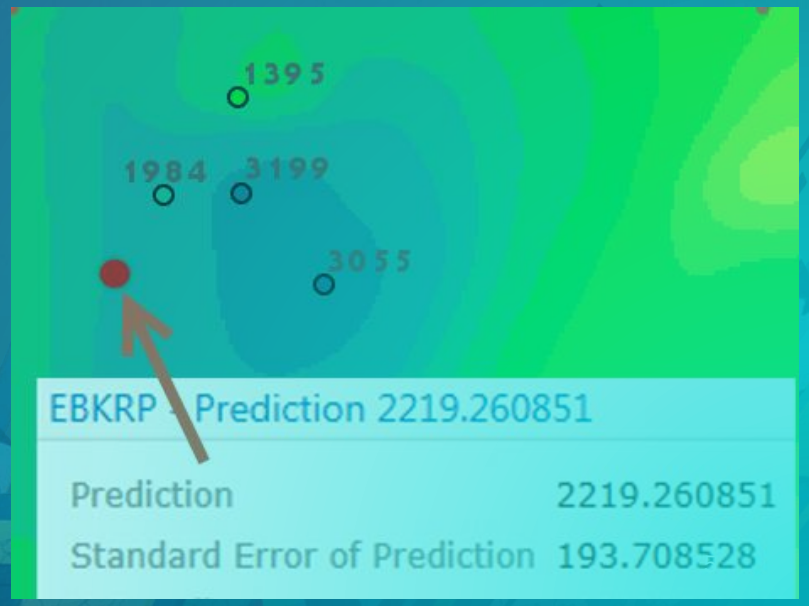
- is a class of statistics used to analyze and predict values associated with spatial phenomena.
- it incorporates the spatial coordinates of the data
- Has evolved to not only provide
  - interpolated values, but also
  - measures of uncertainty

ArcGIS for Geostatistical Analyst

Too expensive to measure everywhere, however, we want to know values everywhere.

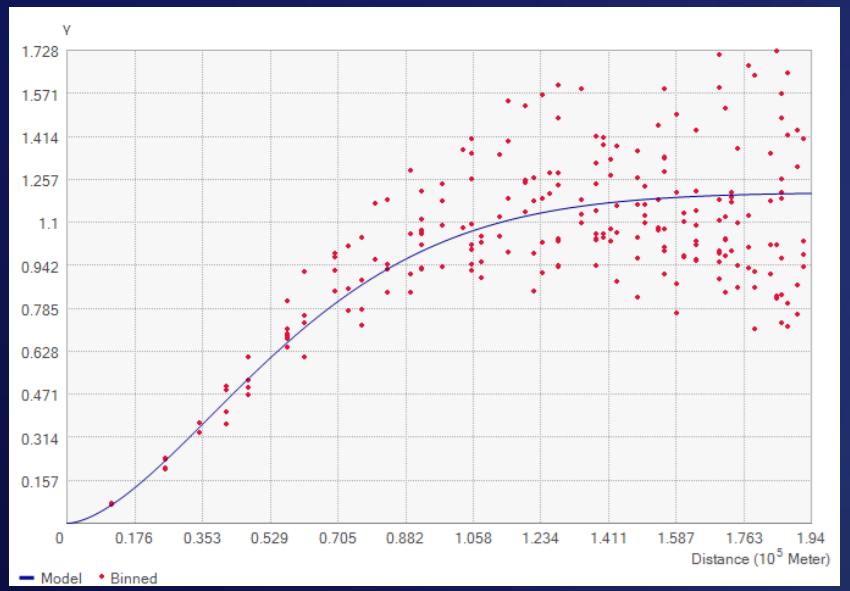


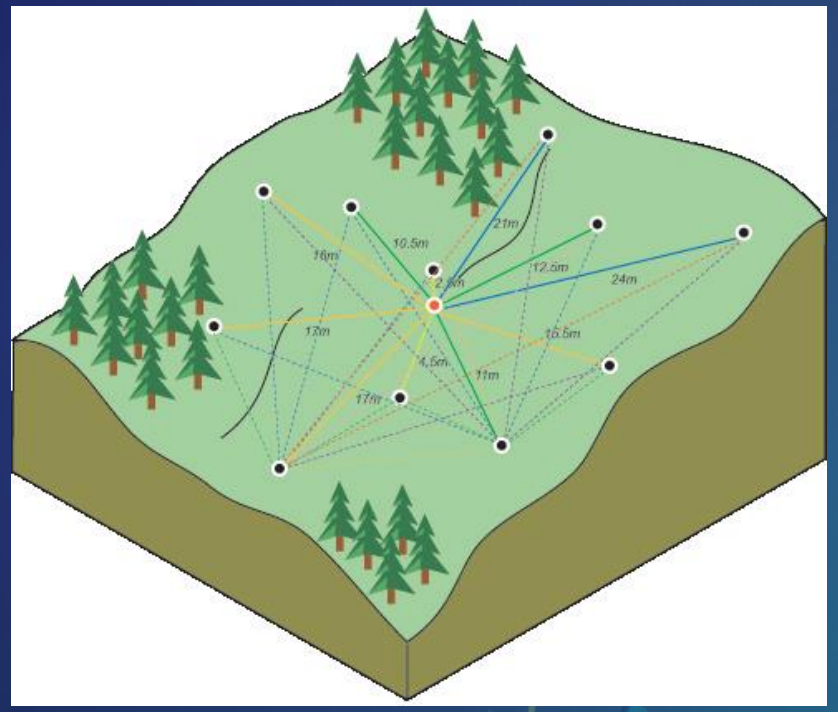


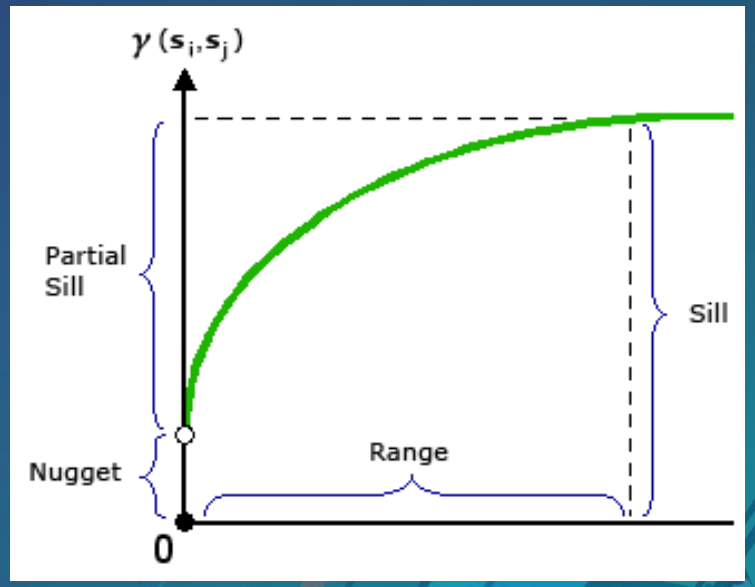


What is a semivariogram?

Semivariogram(distance h) = 0.5 * average [$(\text{value}_i - \text{value}_j)^2$]

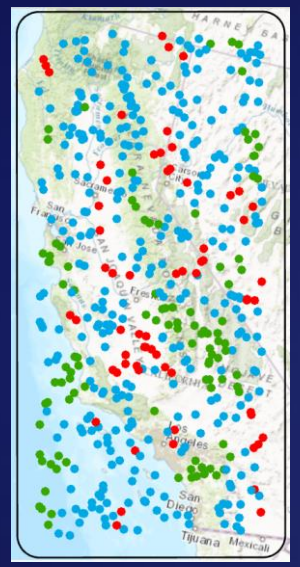


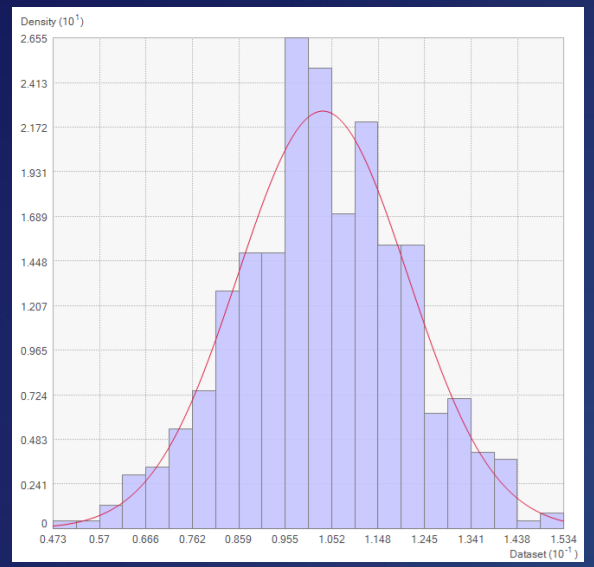


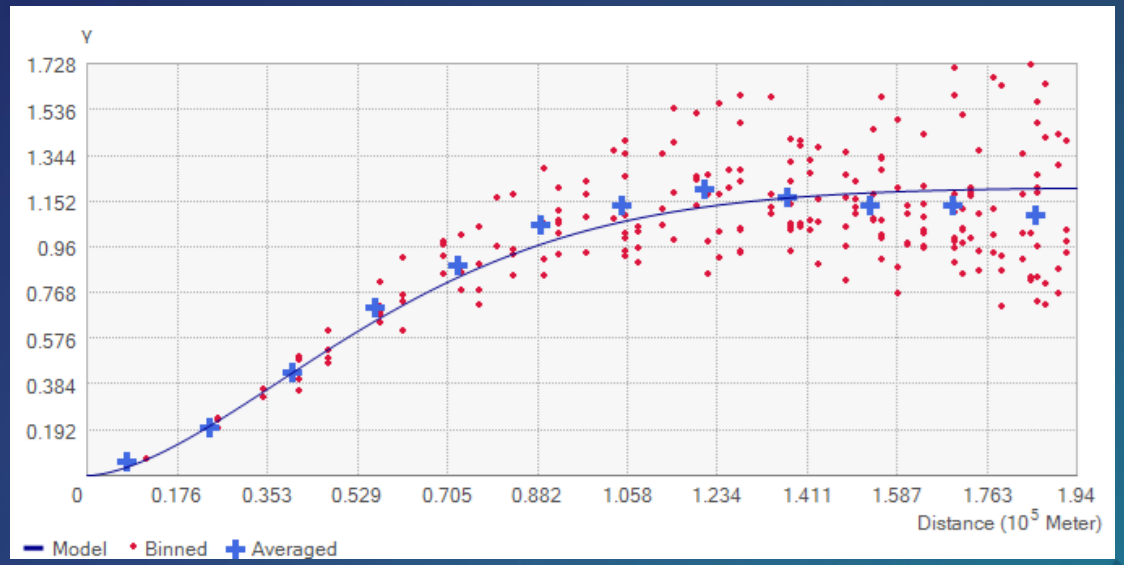


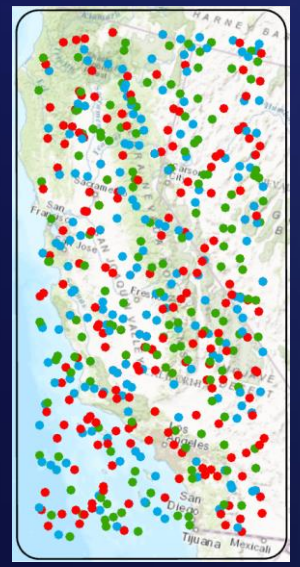
- Range = separation distance between pairs
- Sill = plateau the variogram reaches at the range
- Nugget = sampling error and short scale variability

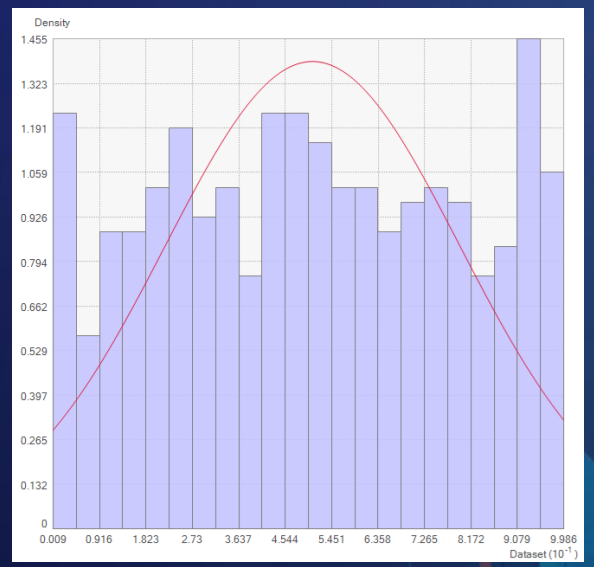
Spatial autocorrelation

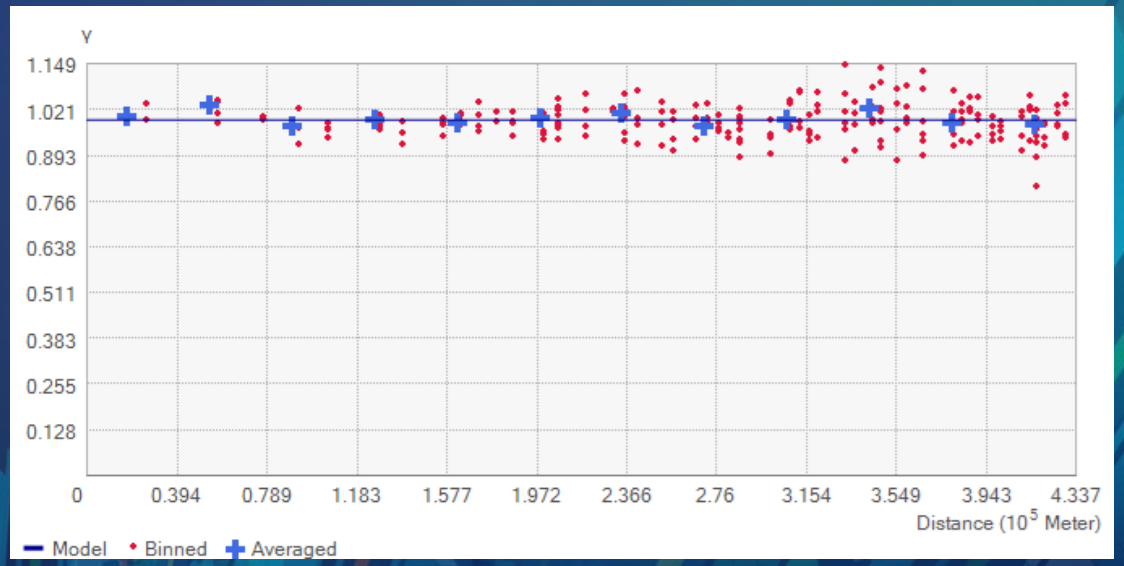








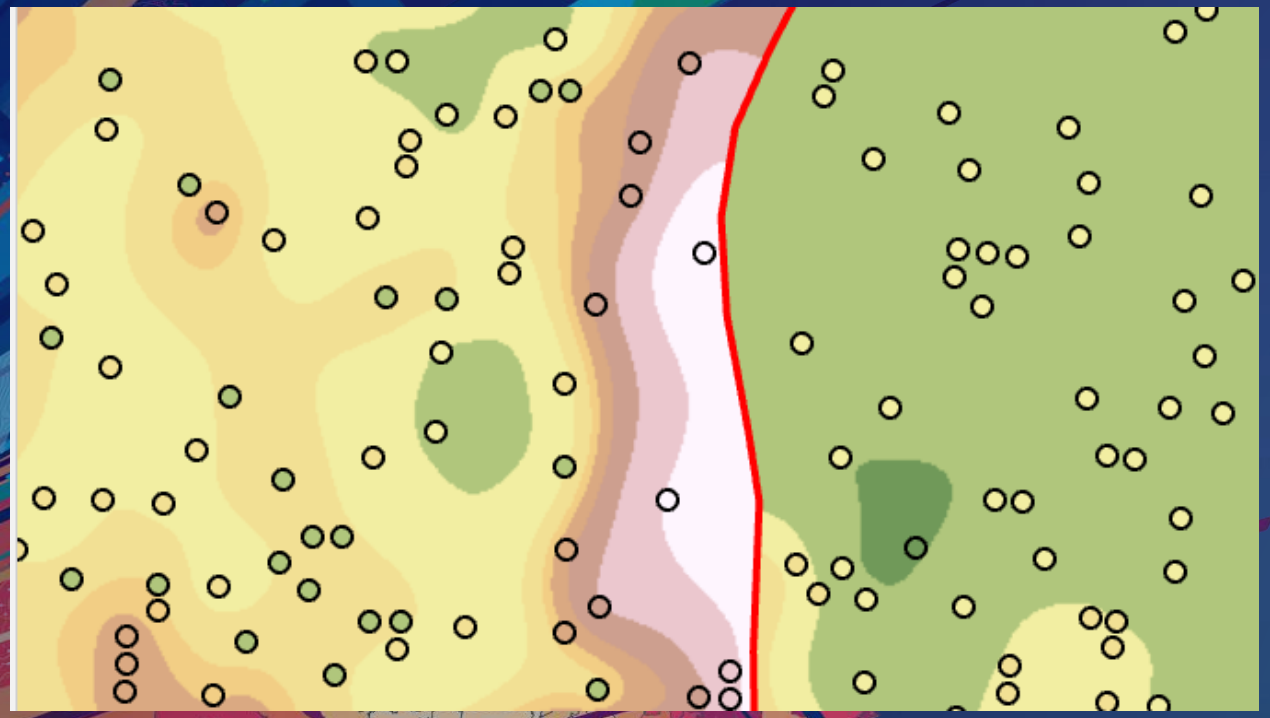




Geostatistical Analyst – What is it?

Provides a complete set of spatial analytical tools that range from techniques to explore the original data to post-processing evaluation of data and predictions uncertainties.

- Geoprocessing tools
  - Use within ArcMap / Pro / Server
  - Modelbuilder
  - Scripting



GP tool

Kernel Interpolation with Barriers

Eric Krause

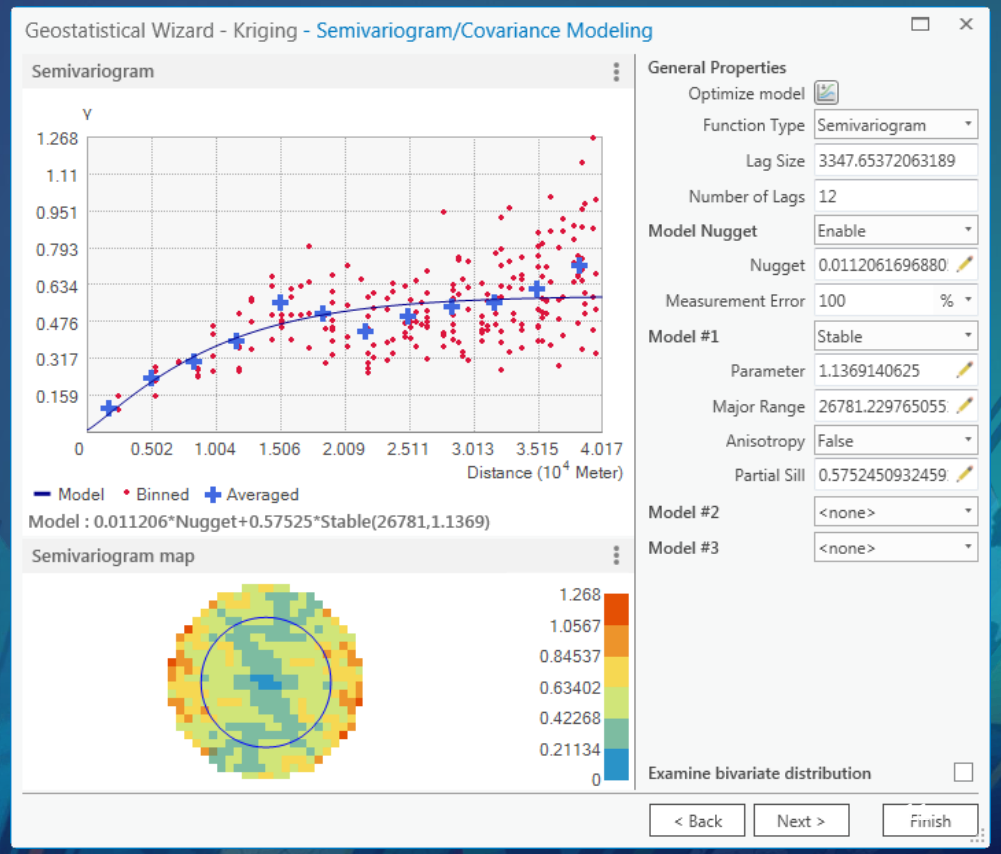
Geostatistical Analyst – Geoprocessing tools

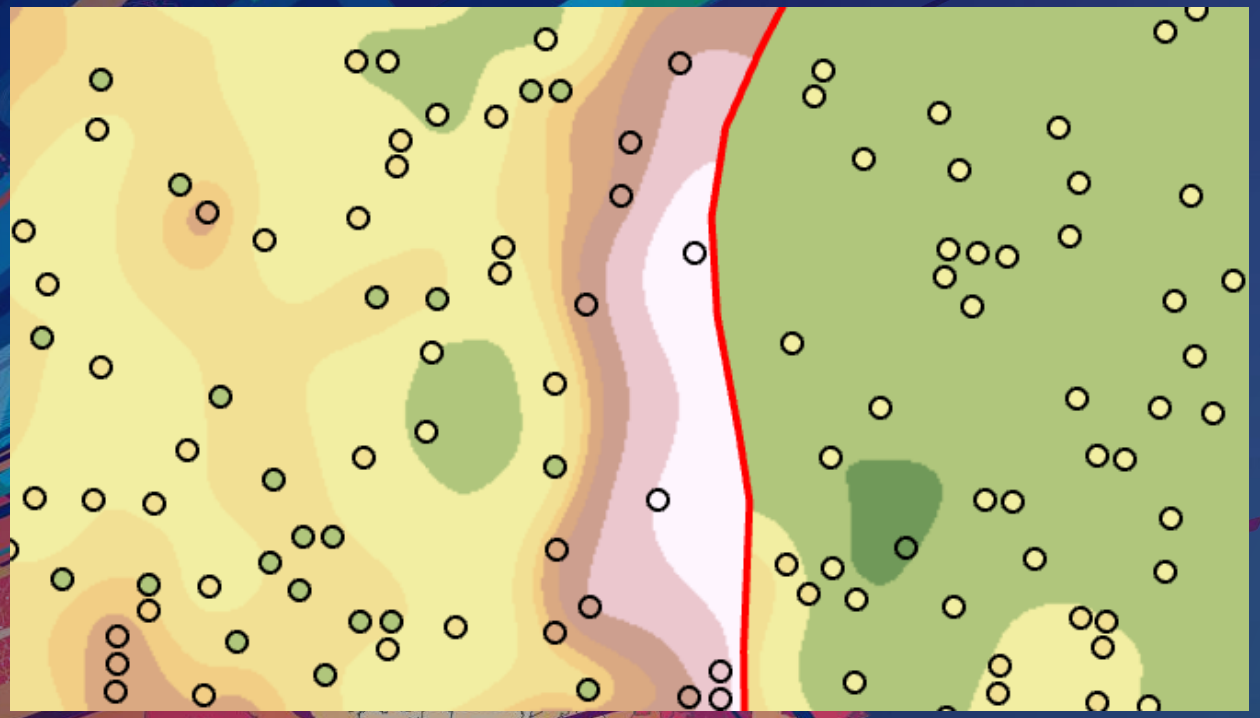
- Geostatistical Analyst Tools
  - Interpolation
    - Diffusion Interpolation With Barriers
    - EBK Regression Prediction
    - Empirical Bayesian Kriging
    - Global Polynomial Interpolation
    - IDW
    - Kernel Interpolation With Barriers
    - Local Polynomial Interpolation
    - Moving Window Kriging
    - Radial Basis Functions
  - Sampling Network Design
    - Create Spatially Balanced Points
    - Densify Sampling Network
  - Simulation
    - Extract Values To Table
    - Gaussian Geostatistical Simulations
  - Utilities
    - Cross Validation
    - Neighborhood Selection
    - Semivariogram Sensitivity
    - Subset Features
  - Working with Geostatistical Layers
    - Areal Interpolation Layer To Polygons
    - Calculate Z-value
    - Create Geostatistical Layer
    - GA Layer To Contour
    - GA Layer To Grid
    - GA Layer To Points
    - GA Layer To Rasters
    - Get Model Parameter
    - Set Model Parameter

Geostatistical Analyst – What is it?

Provides a complete set of spatial analytical tools that range from techniques to explore the original data to post-processing evaluation of data and predictions uncertainties.

- Wizard
  - is a dynamic set of pages that is designed to guide you through the process of constructing and evaluating the performance of an interpolation model.





Geostatistical Wizard

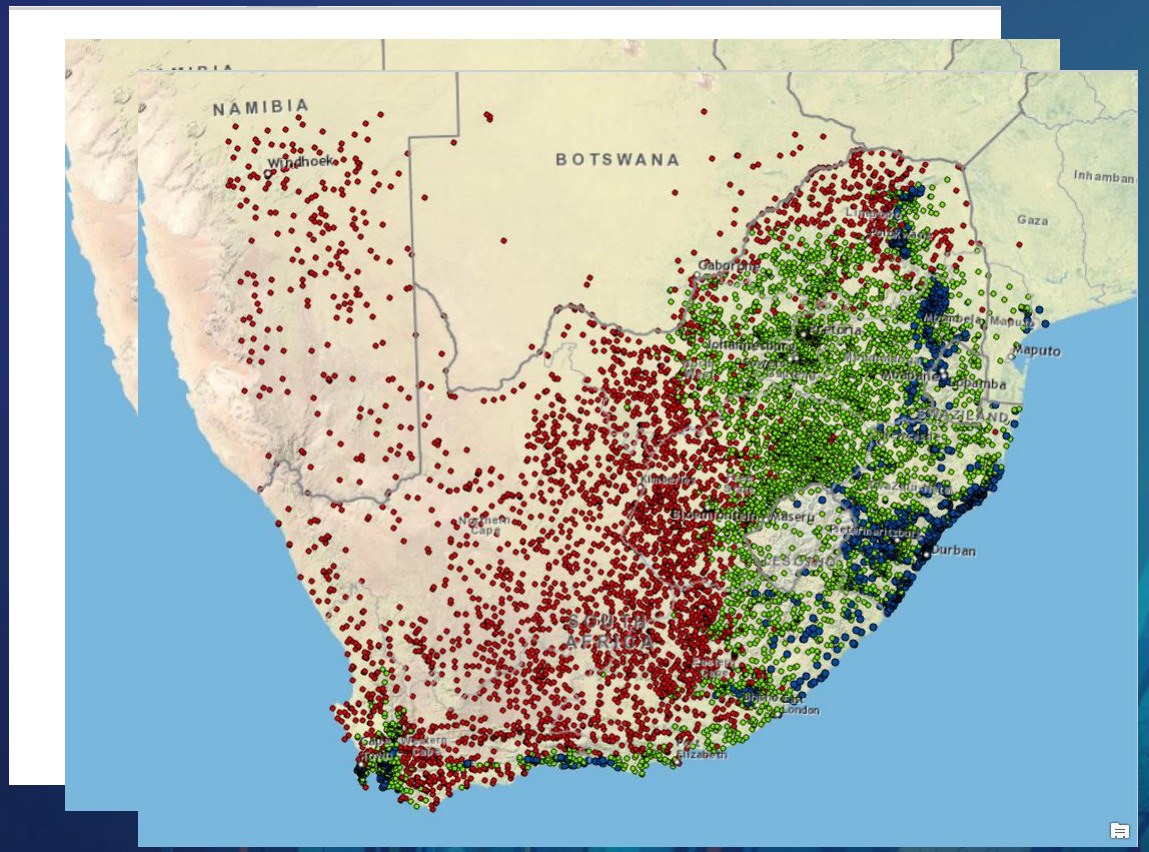
Kernel Interpolation with Barriers

Eric Krause

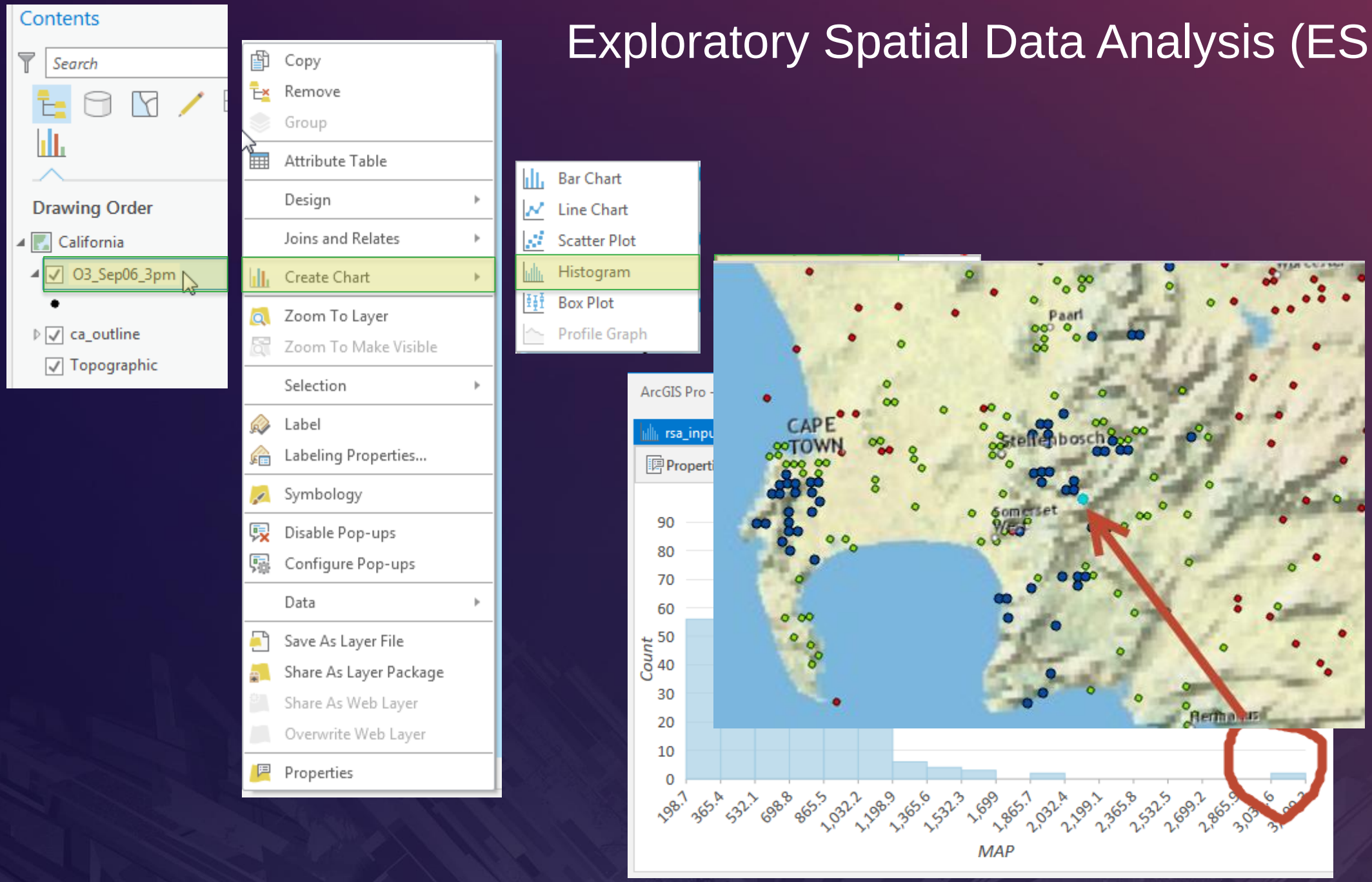
ESDA

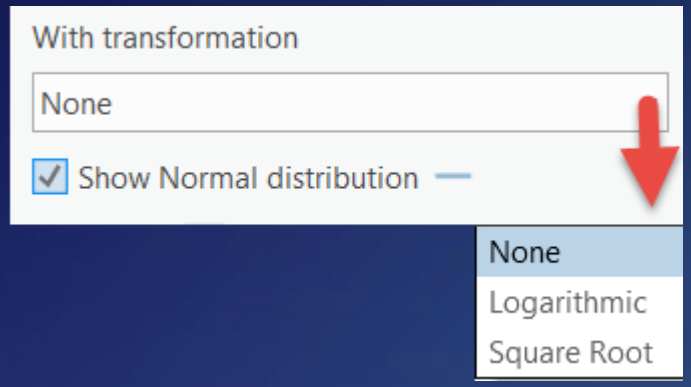
Exploratory Spatial Data Analysis

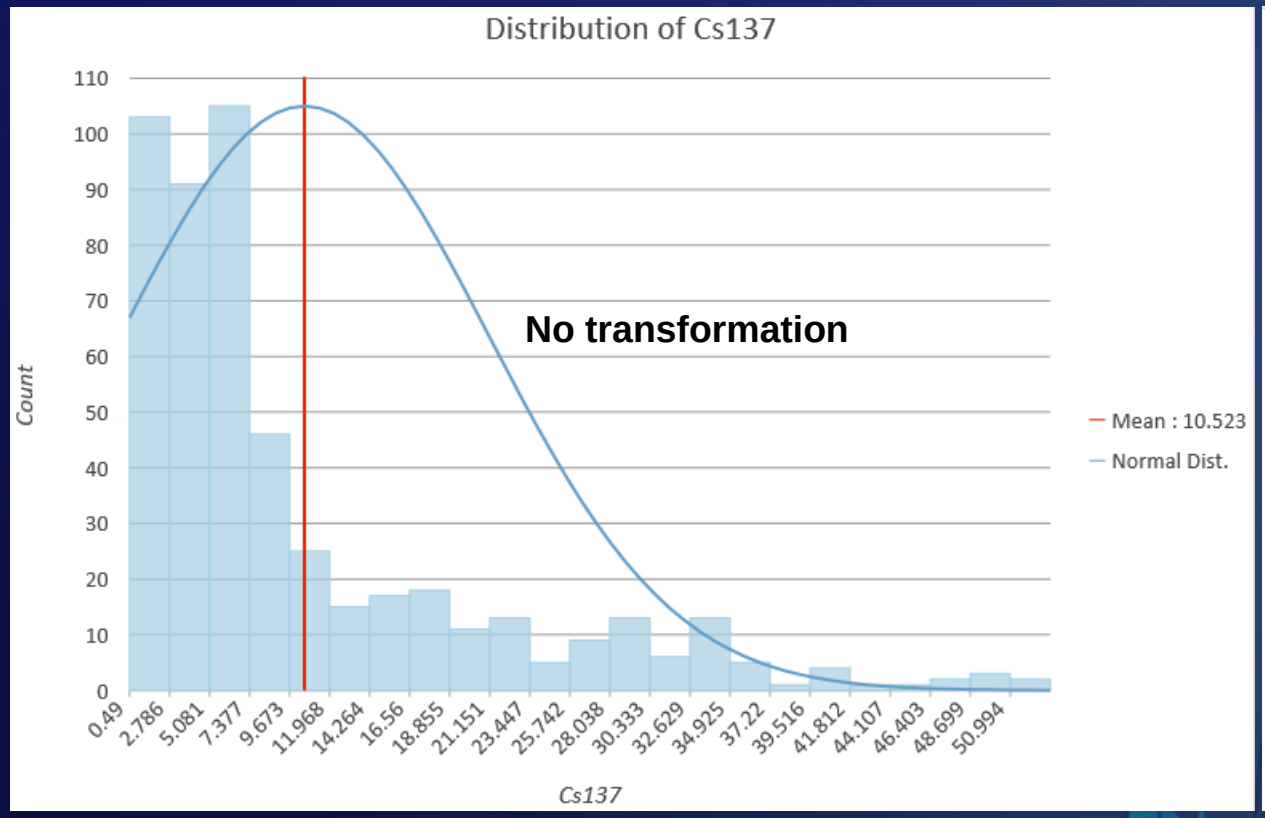
- Where is the data located?
- What are the values at the data points?
- How does the location of a point relate to its value?

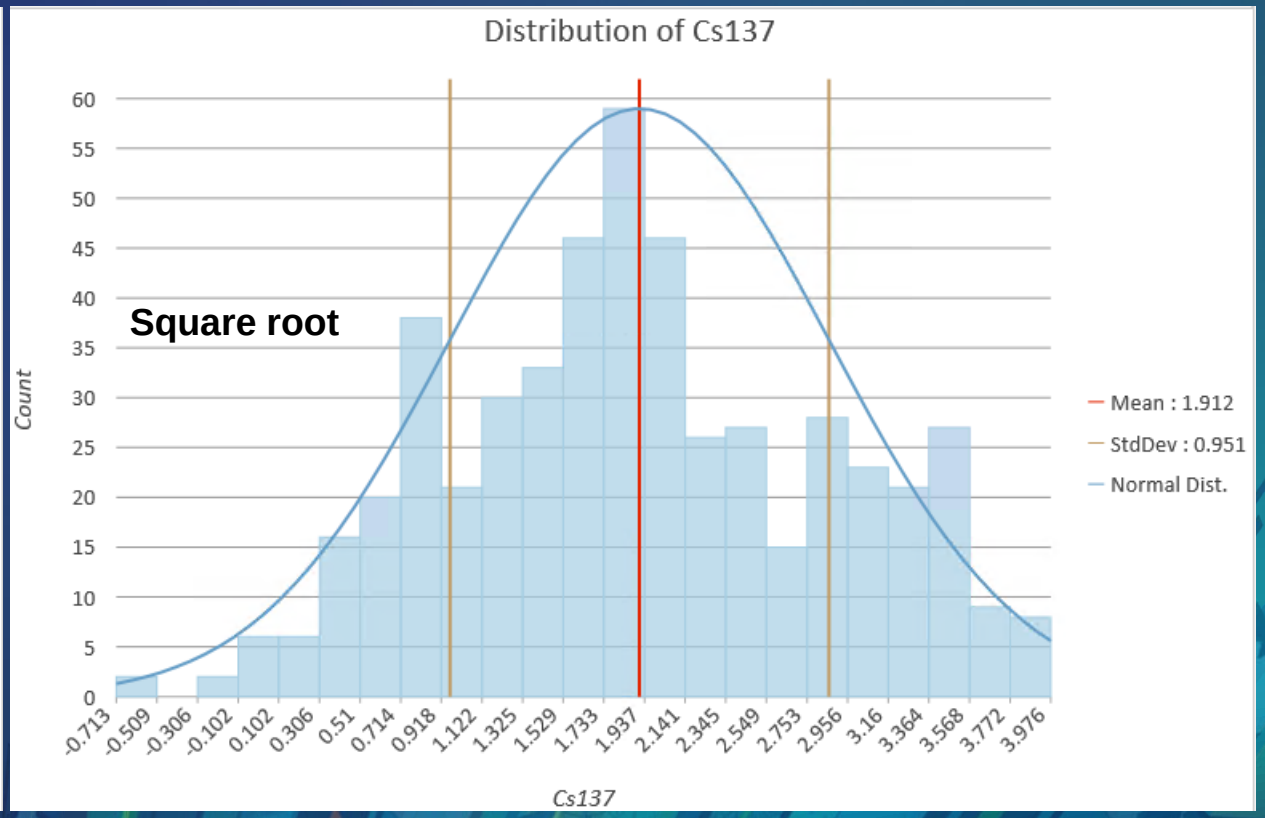


Exploratory Spatial Data Analysis (ESDA)




With transformation
None
Show Normal distribution
None
Logarithmic
Square Root


Distribution of Cs137
No transformation
Count
Cs137
Mean : 10.523
Normal Dist.


Distribution of Cs137
Square root
Count
Cs137
Mean : 1.912
StdDev : 0.951
Normal Dist.

What is kriging?

Tuesday

- Kriging: An Introduction to Concepts and Applications (10:00-11:00 SDCC Rm02)
- Empirical Bayesian Kriging and EBK Regression Prediction in ArcGIS (2:30-3:15 SDCC Demo Theater 12)

Wednesday

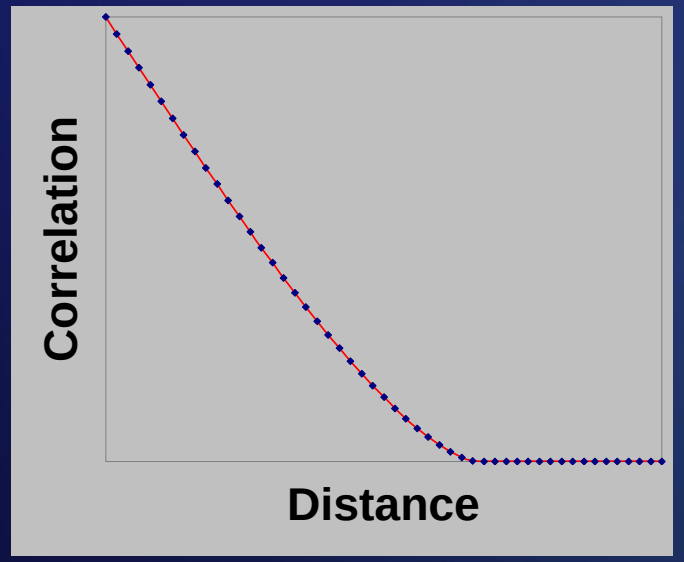
- Kriging: An Introduction to Concepts and Applications (4:00-5:00 SDCC Rm08)

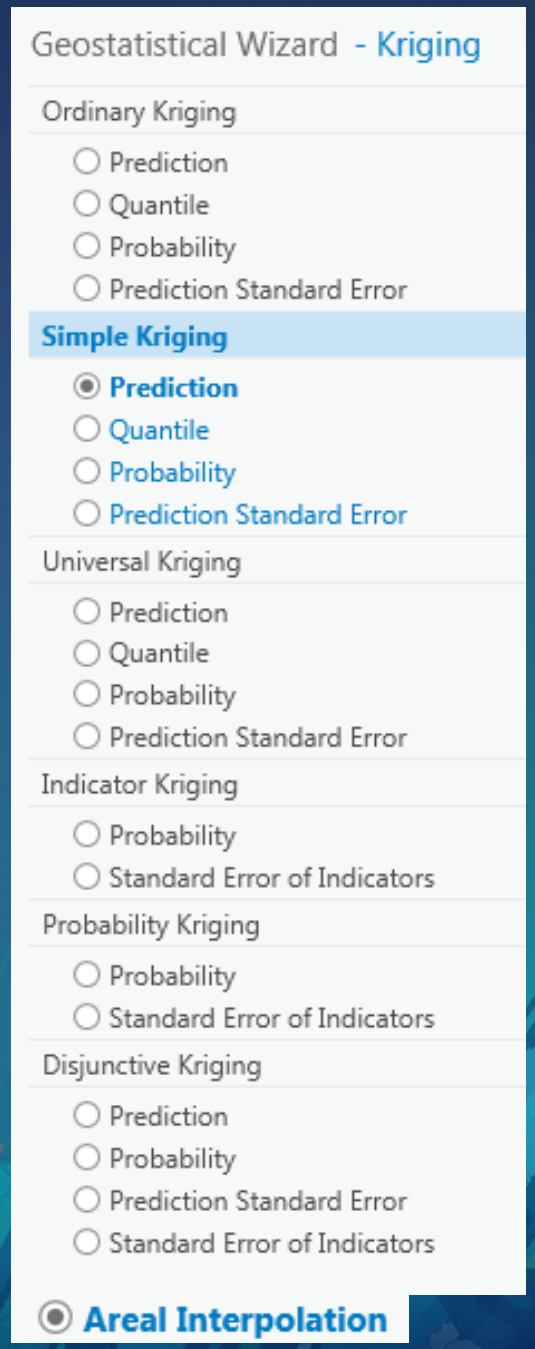
Thursday

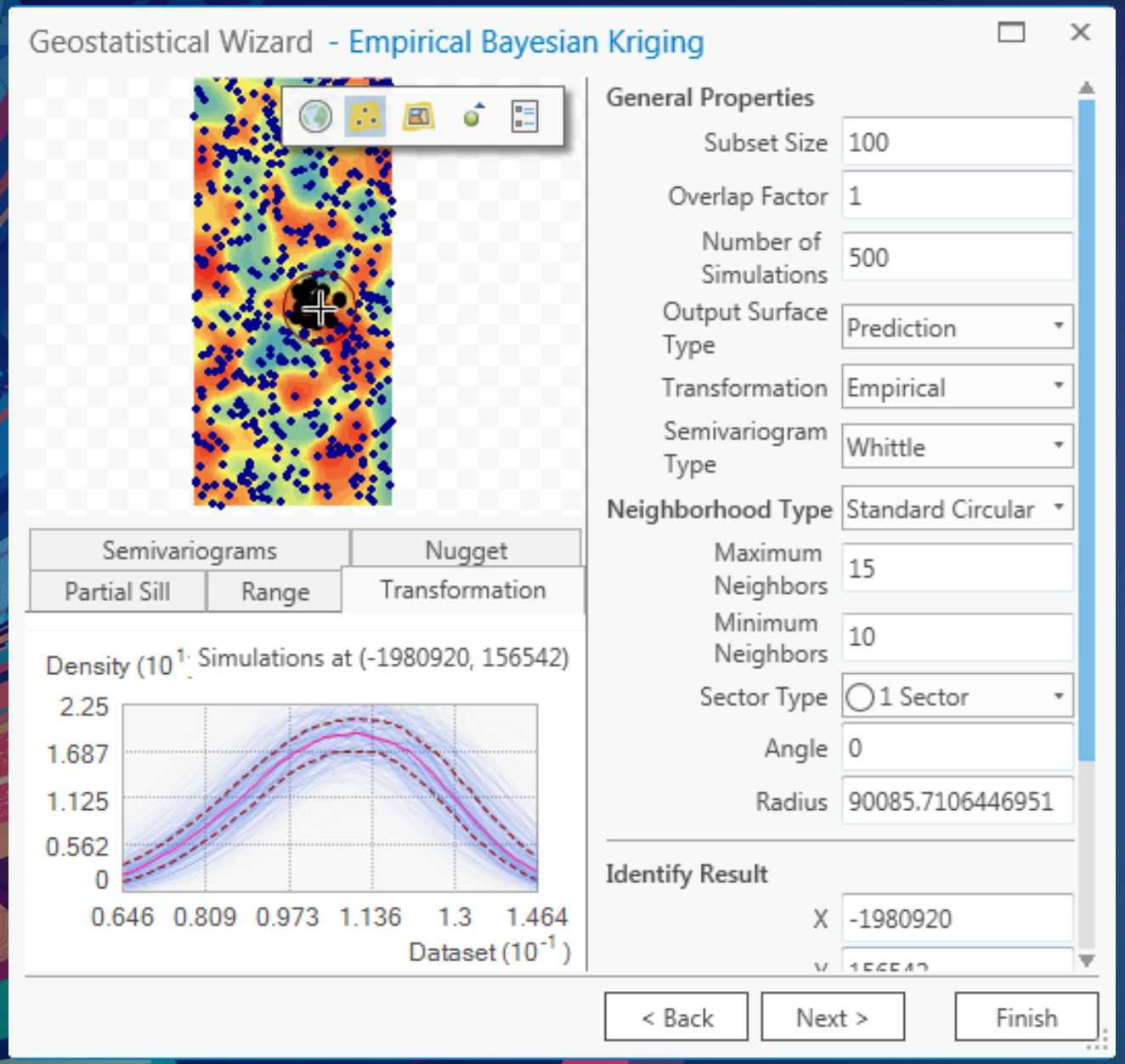
- Polygon-to-Polygon Predictions Using Areal Interpolation (11:15-12:00 SDCC Demo Theater 10)

What is kriging?

- It is a geostatistical interpolation technique
- that models the spatial correlation of point measurements
- to estimate values at unmeasured locations.
- Associates uncertainty with the predictions







Geostatistical Wizard

Eric Krause

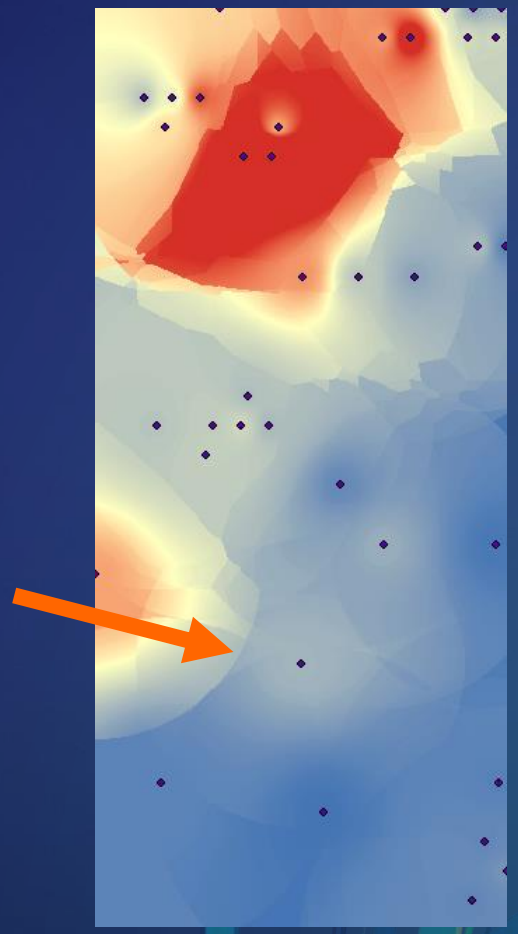
Interpolation workflow

- ESDA
- Interpolate
- Goodness of fit

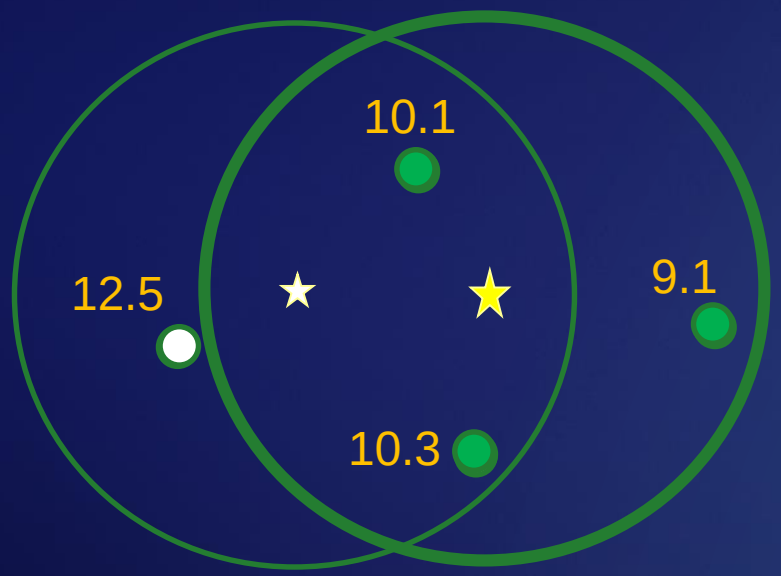
Why use ESRI's Geostatistical Analyst?

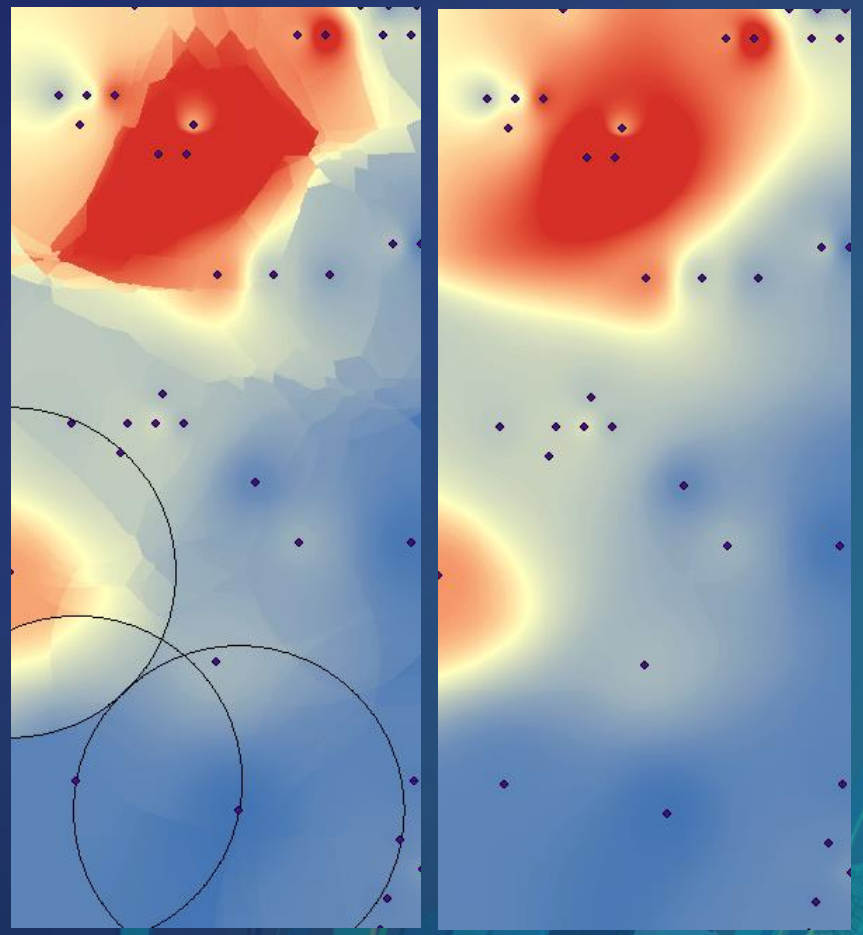
- Search neighborhood
  - Sectors
  - Smooth
- Chordal distance
- Cross validation
- Error maps
- Interactive Variography
- Barriers
- Simulations

Search neighborhood - Smooth



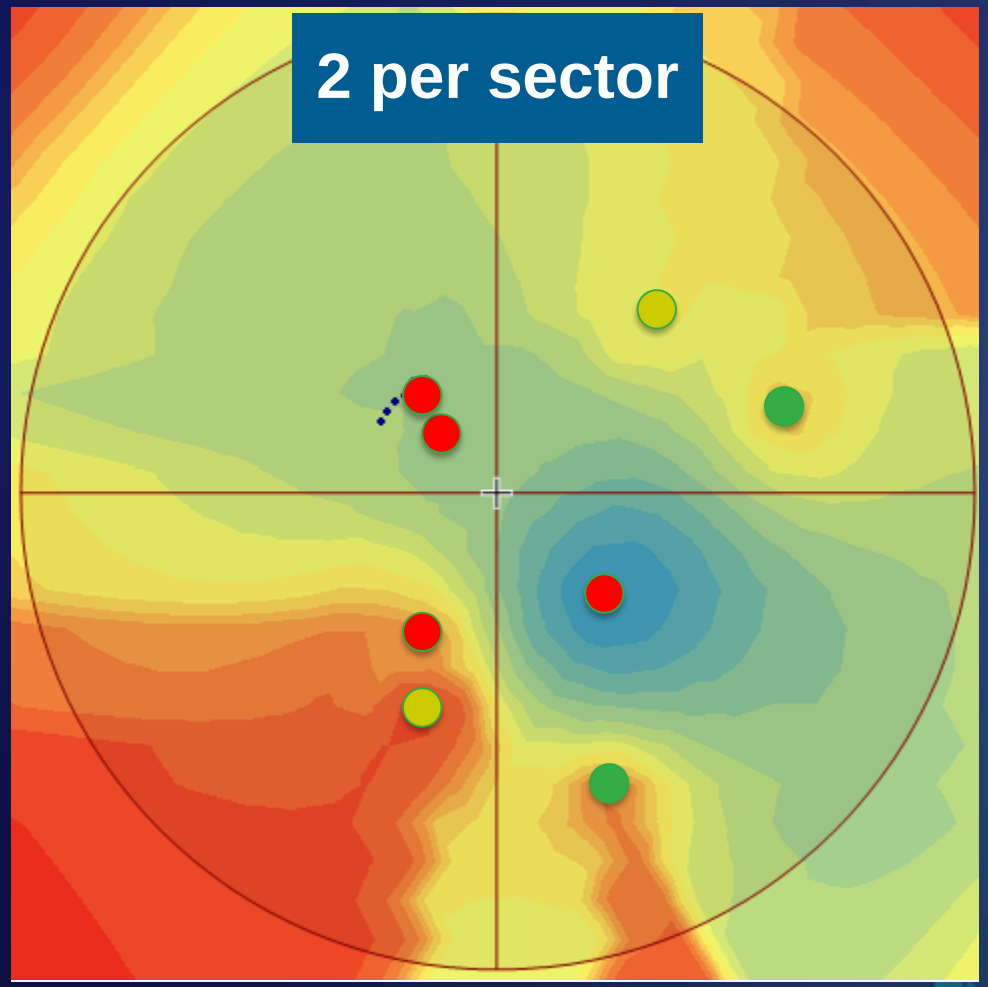
Search neighborhood - Smooth

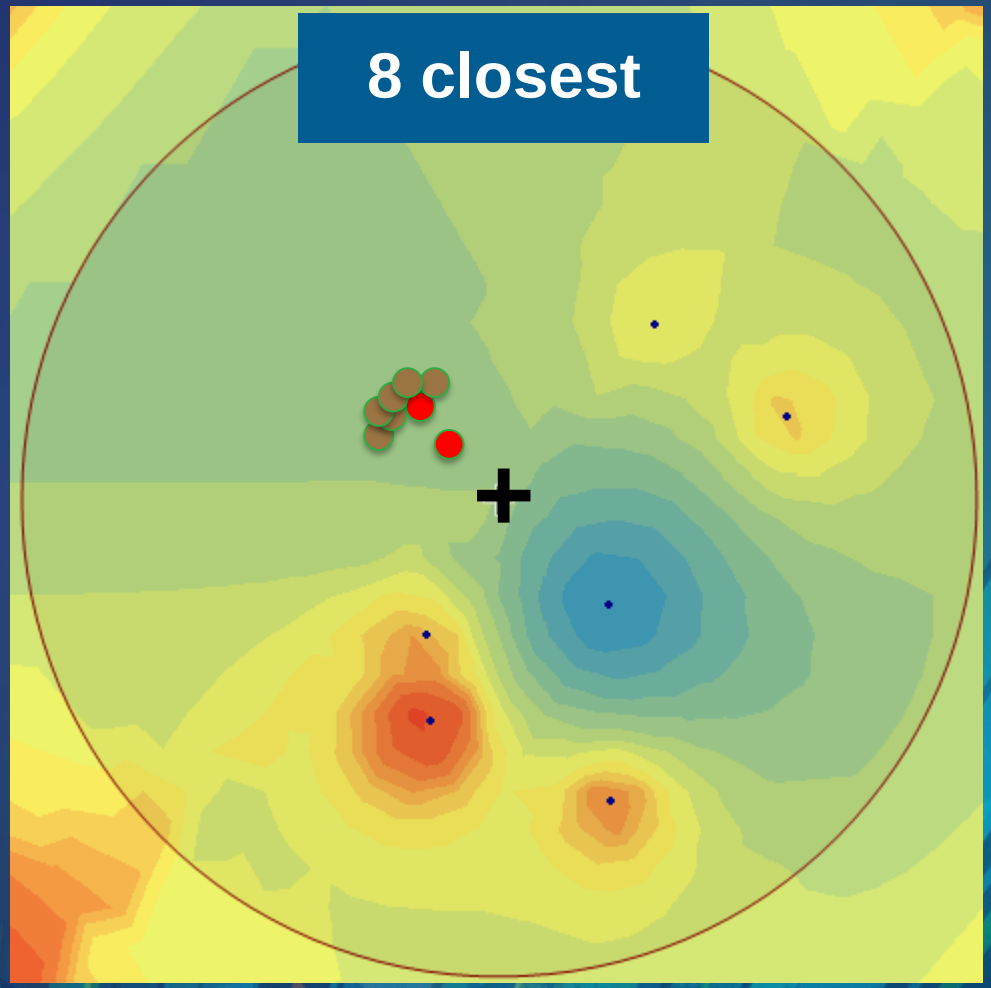




Unlike smoothing the output,
this method modifies the weights

Search neighborhood - Standard





Chordal distances

Only for EBK and EBK Regression Prediction

- Automatically kicks in when data are in GCS
- The chordal distance between any two points is the straight-line distance that connects the two points.
- This line will go through the earth rather than along its surface.

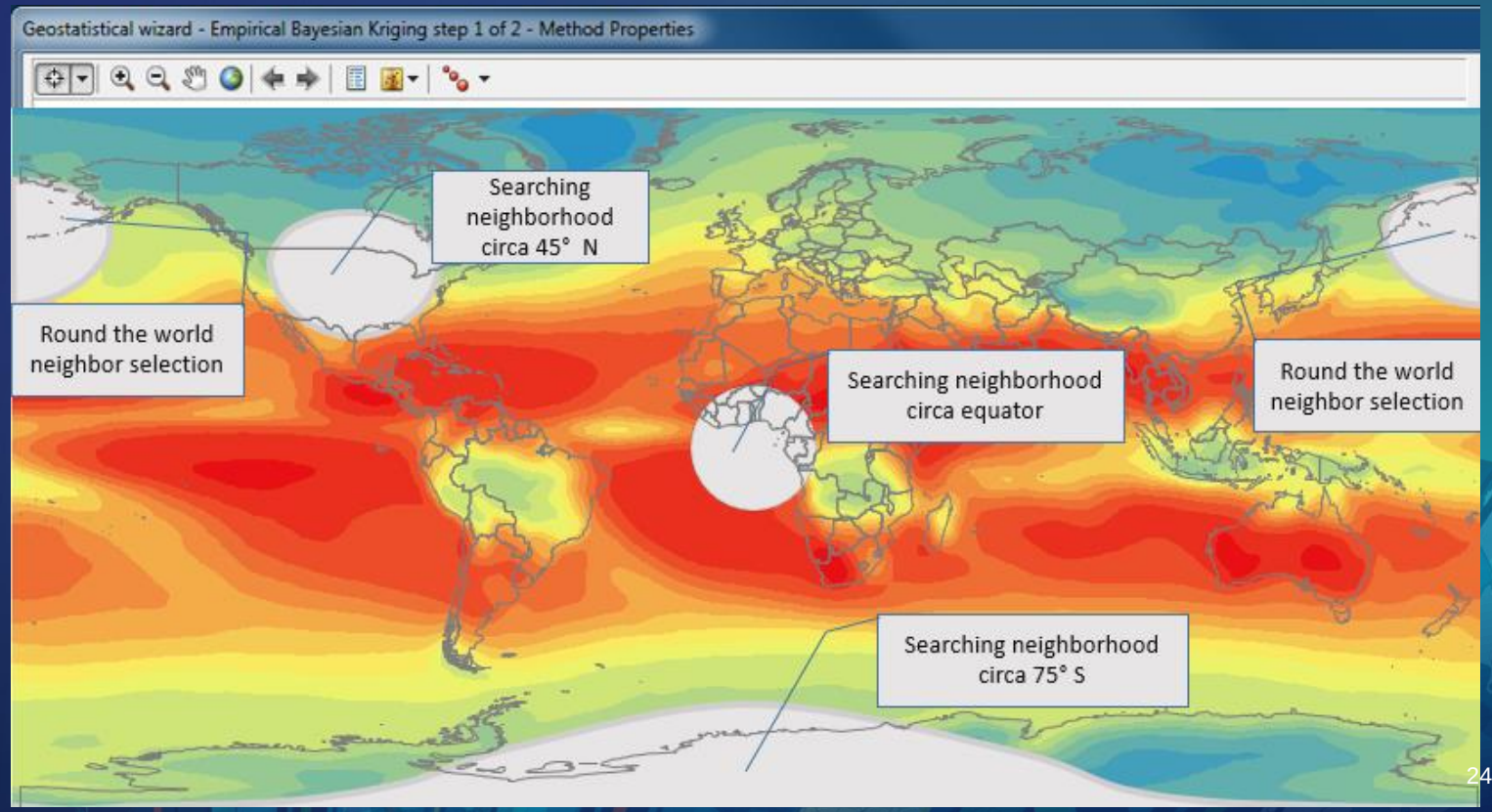
Distance between LA and New York

Geodesic = 3,939.1 km

Chordal = 3,877.0 km

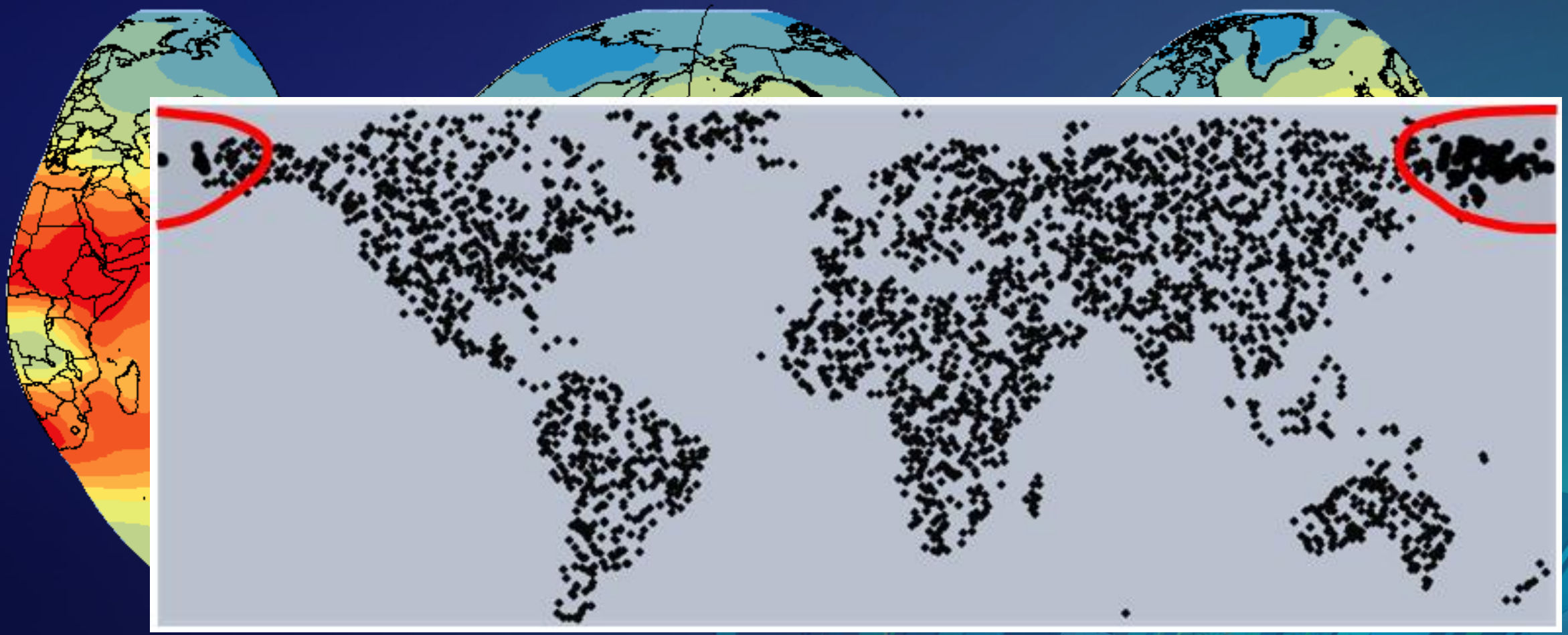
Difference = 62.1 km (1.5%)

Speed!

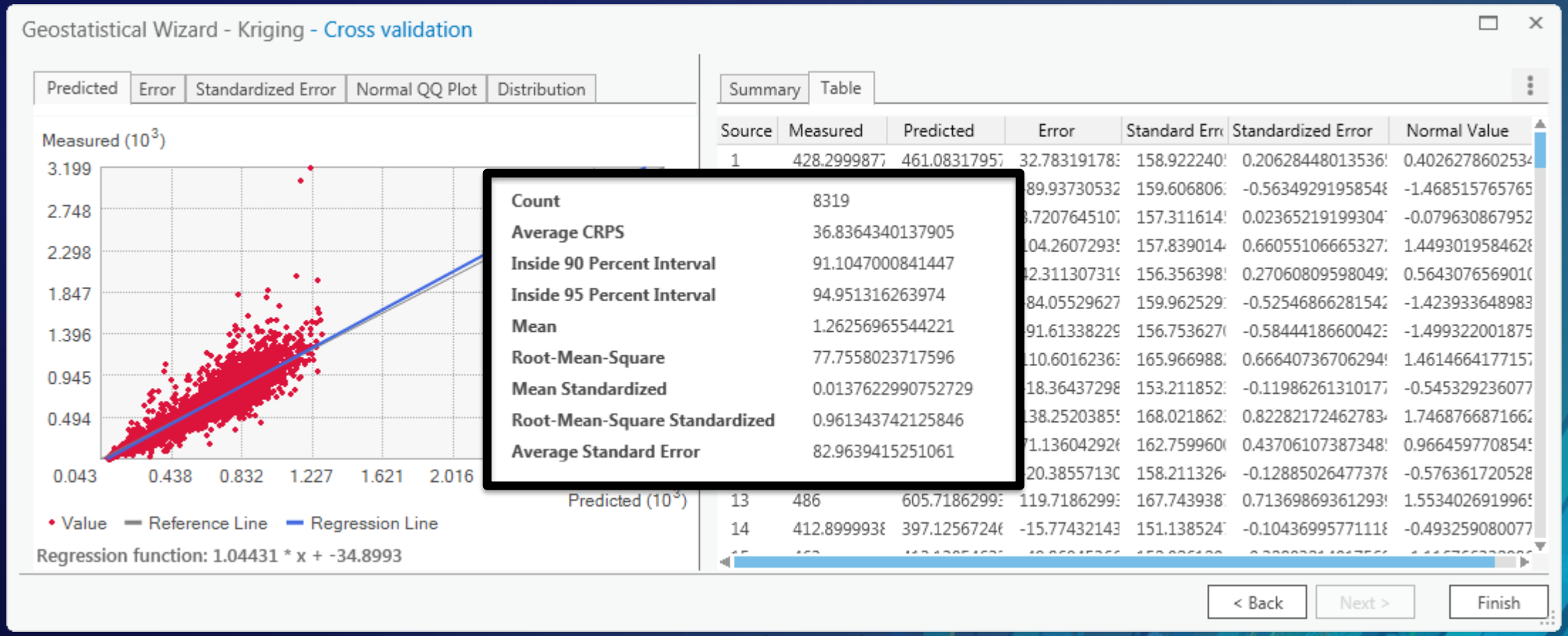


Chordal distances

Only for EBK and EBK Regression Prediction

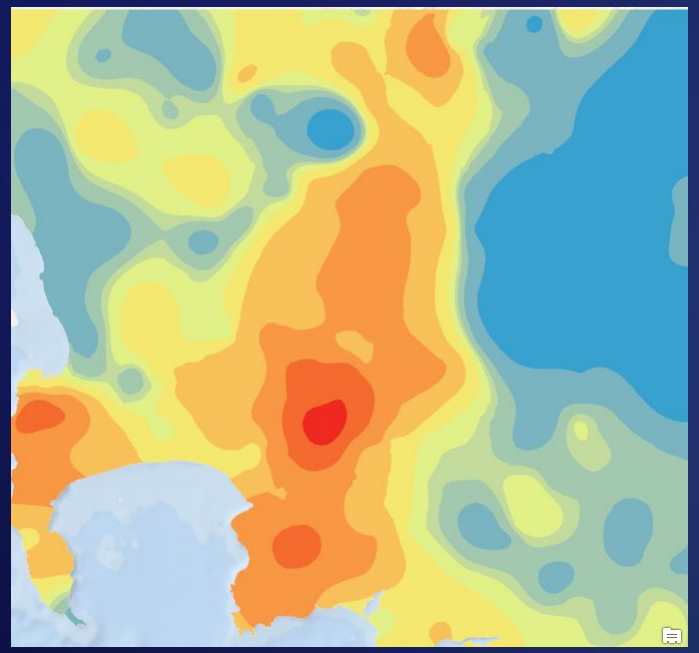


Cross validation / Validation

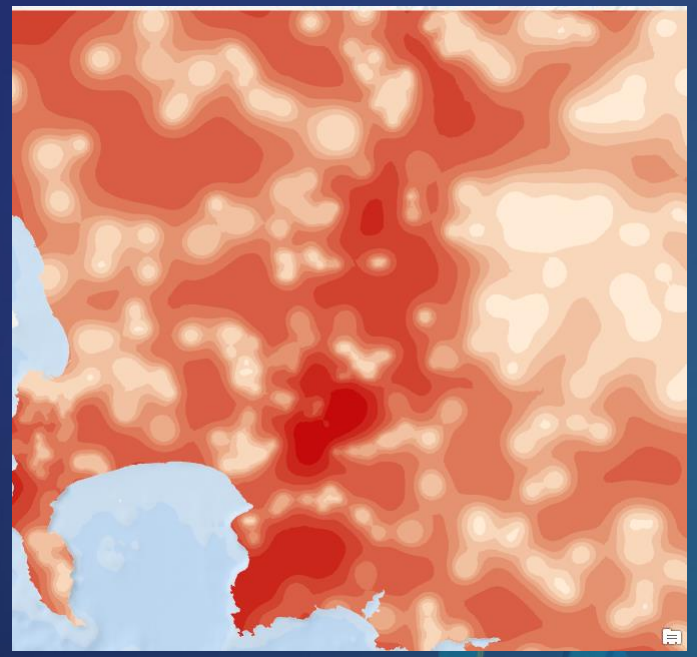


Output surfaces

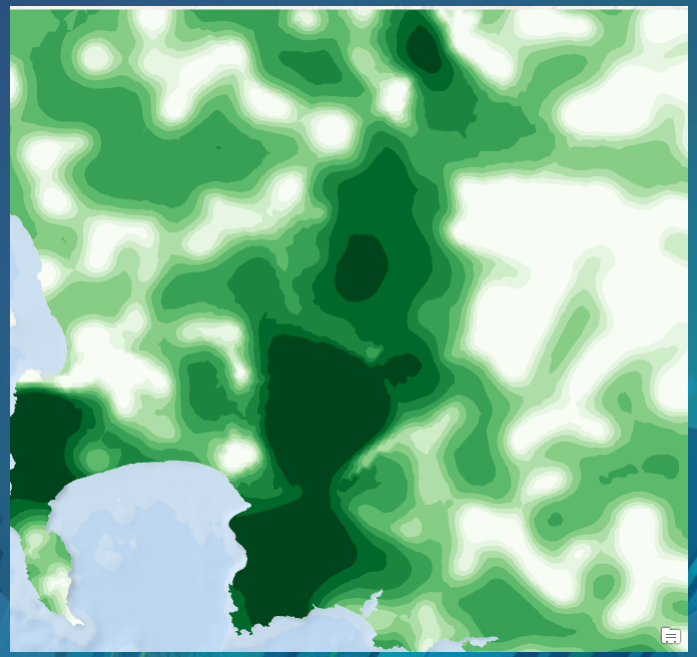
Prediction



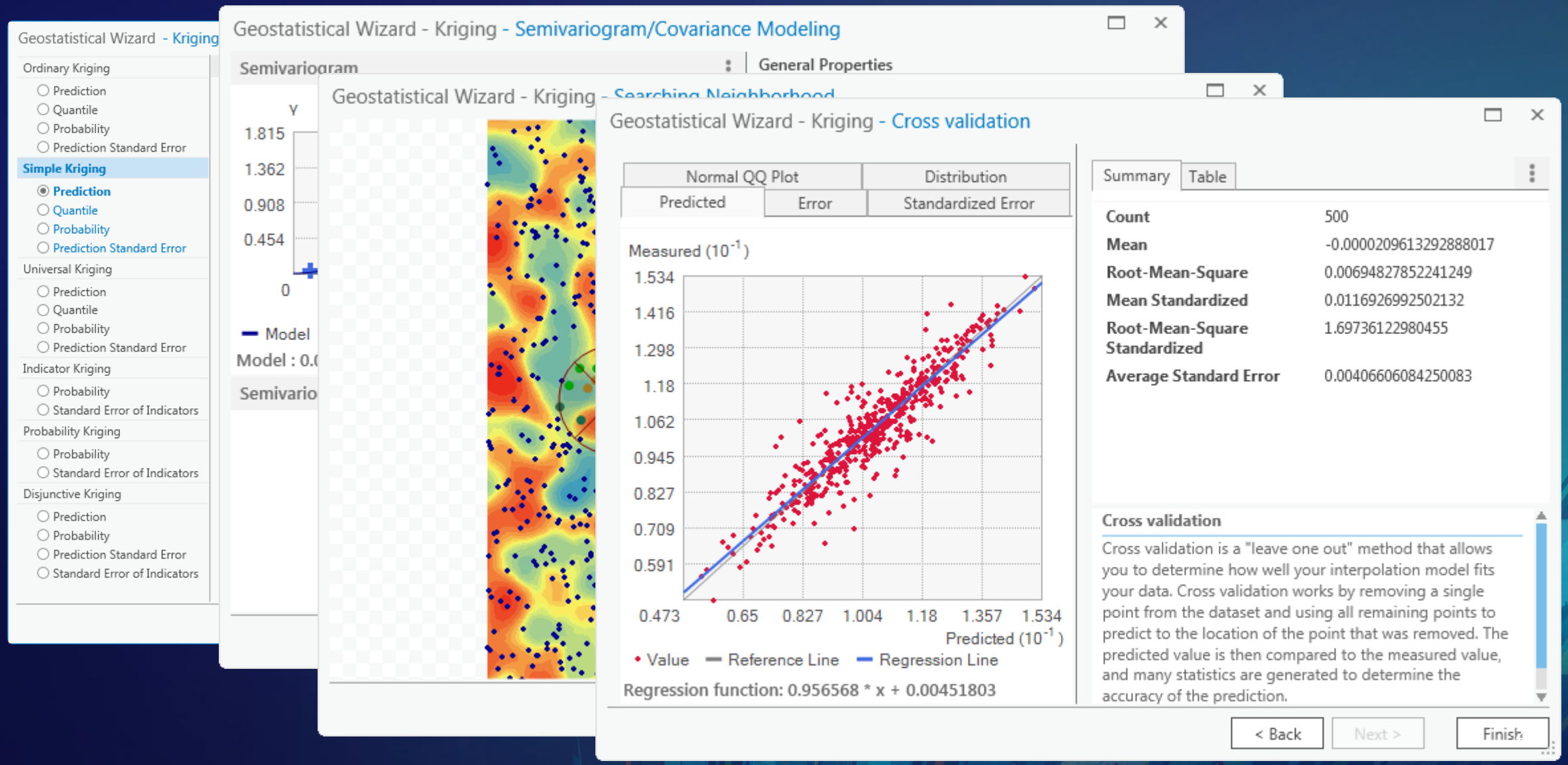
Standard error of prediction



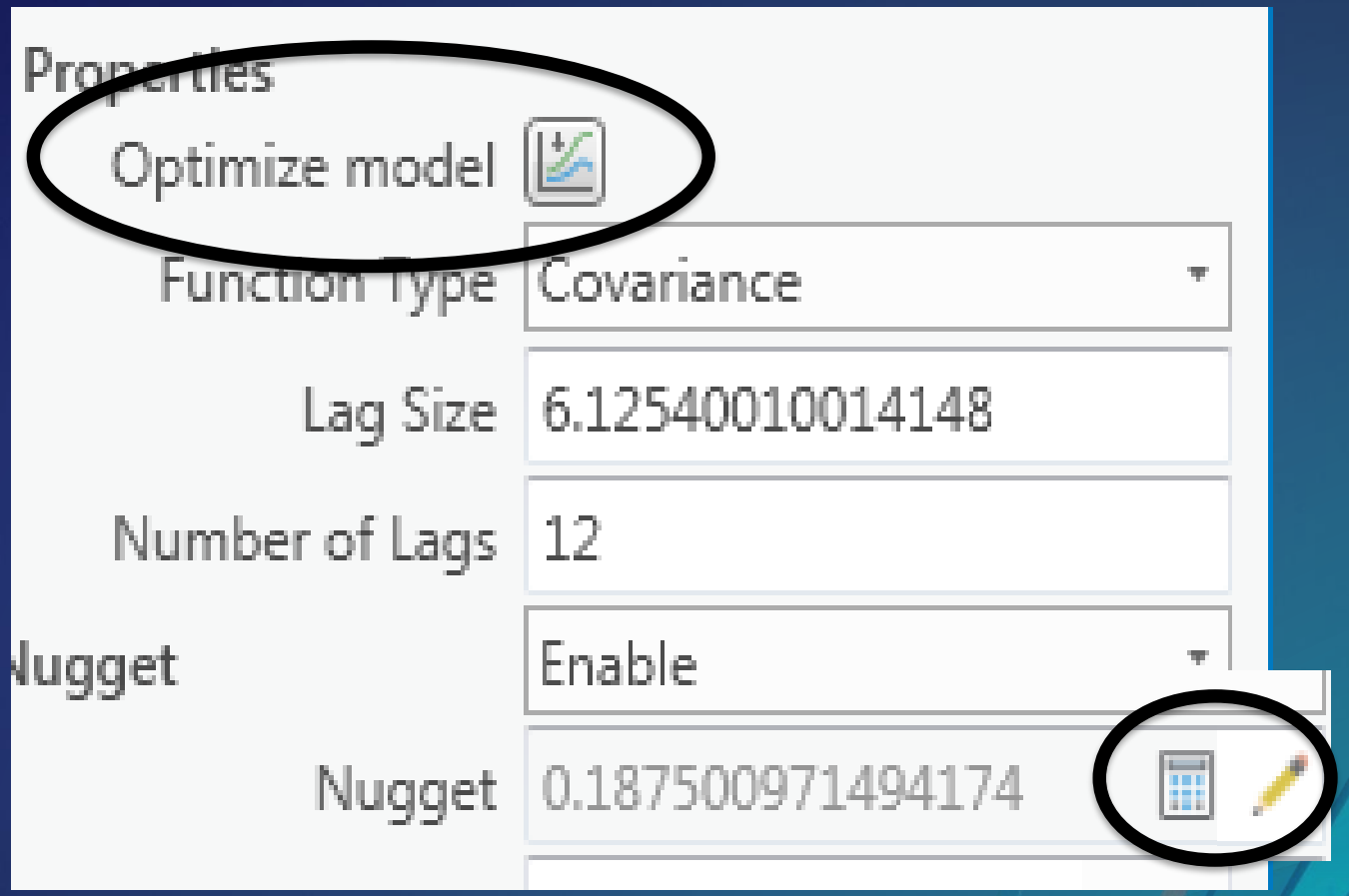
Probability that rainfall exceeds 900mm



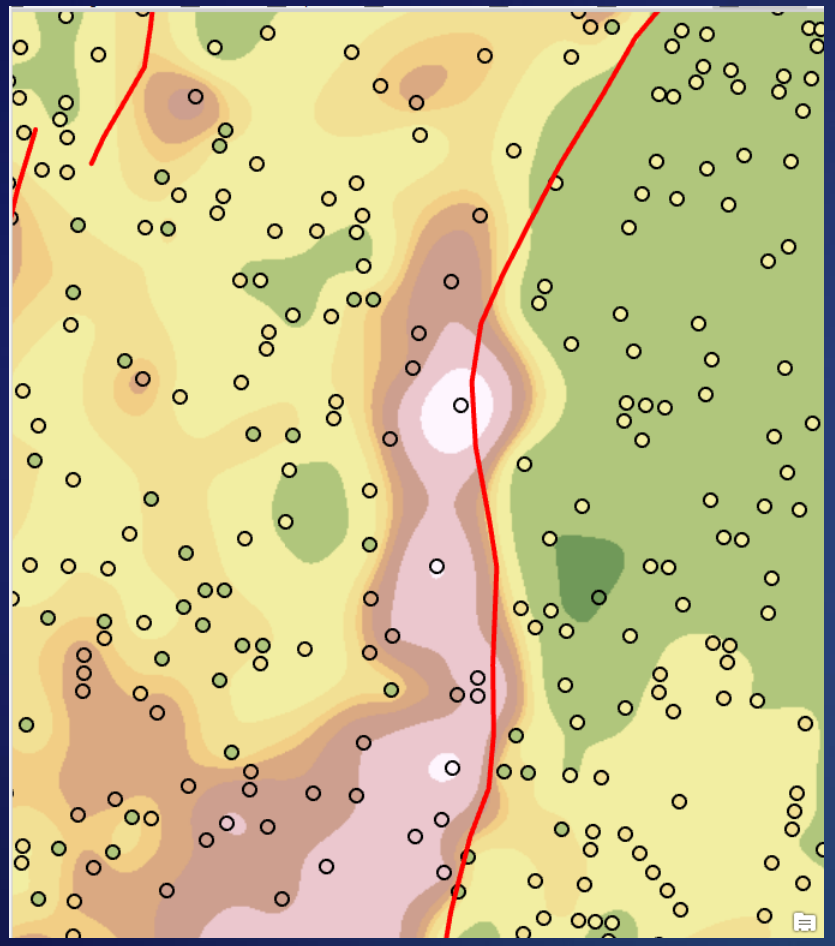
Interactive Wizard

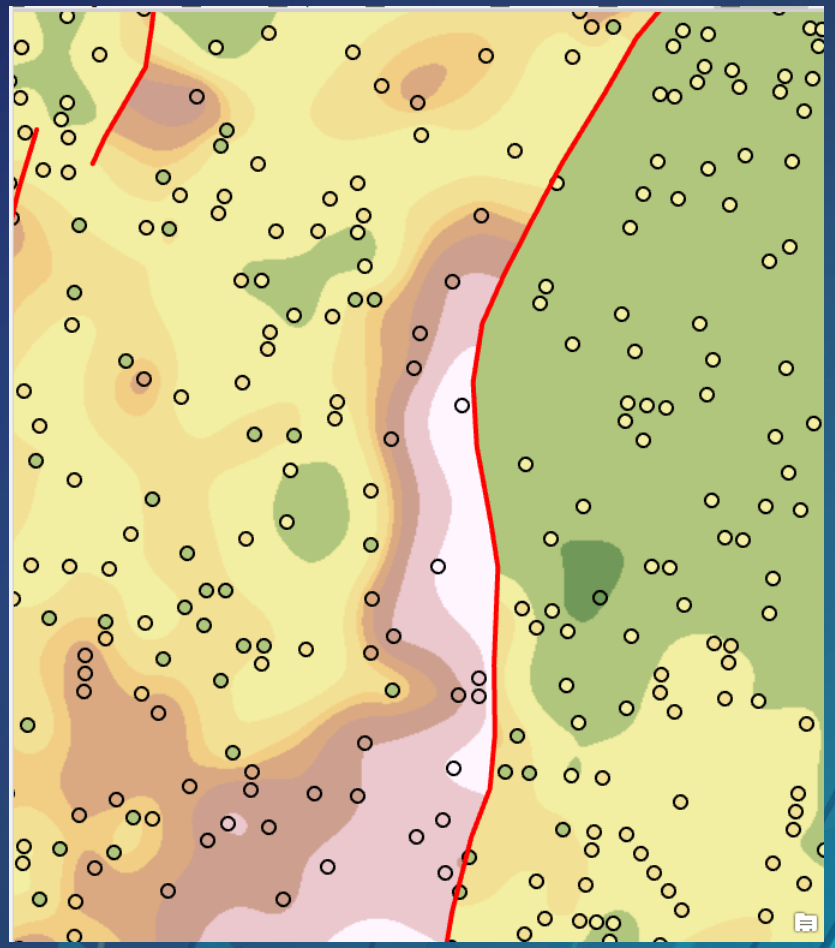


Interactive Wizard



Barriers

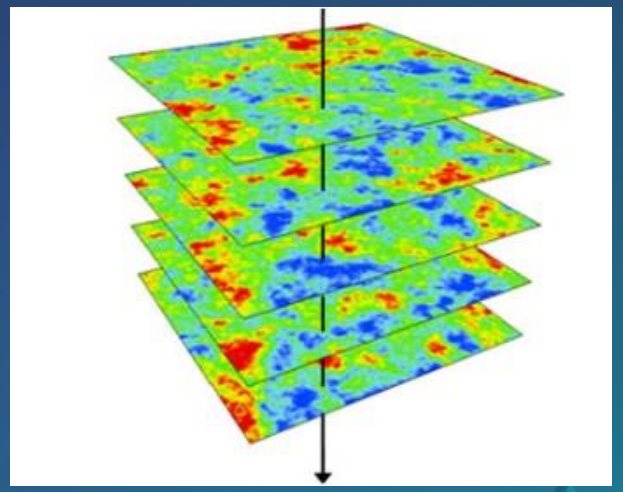


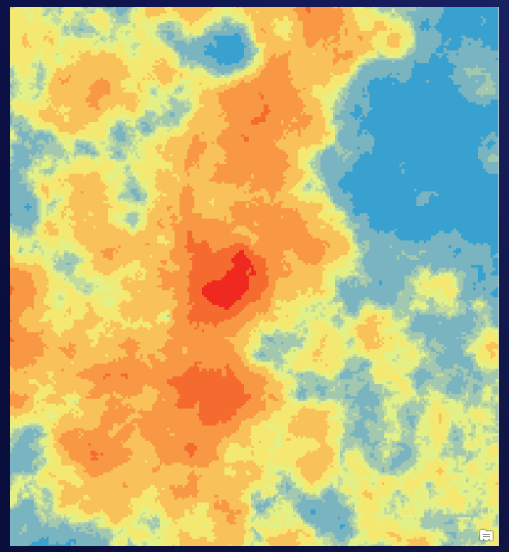


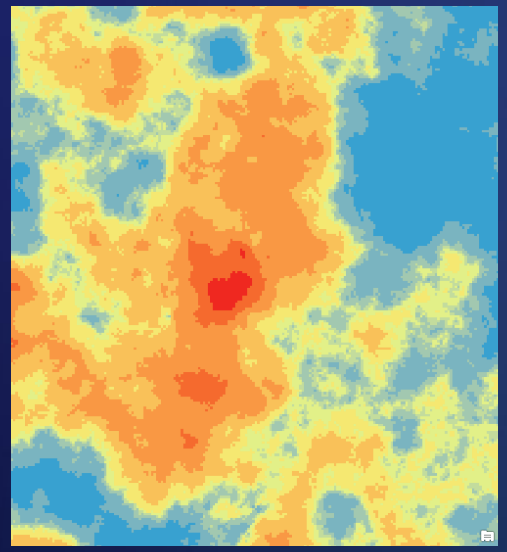
Gaussian Geostatistical Simulations

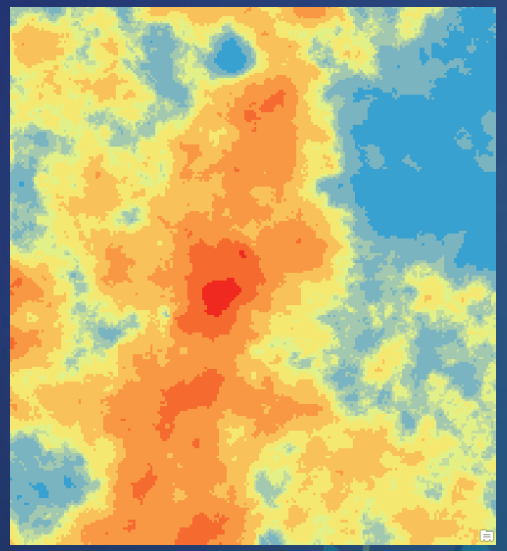
Create multiple versions (realizations) of a surface to perform risk analysis.

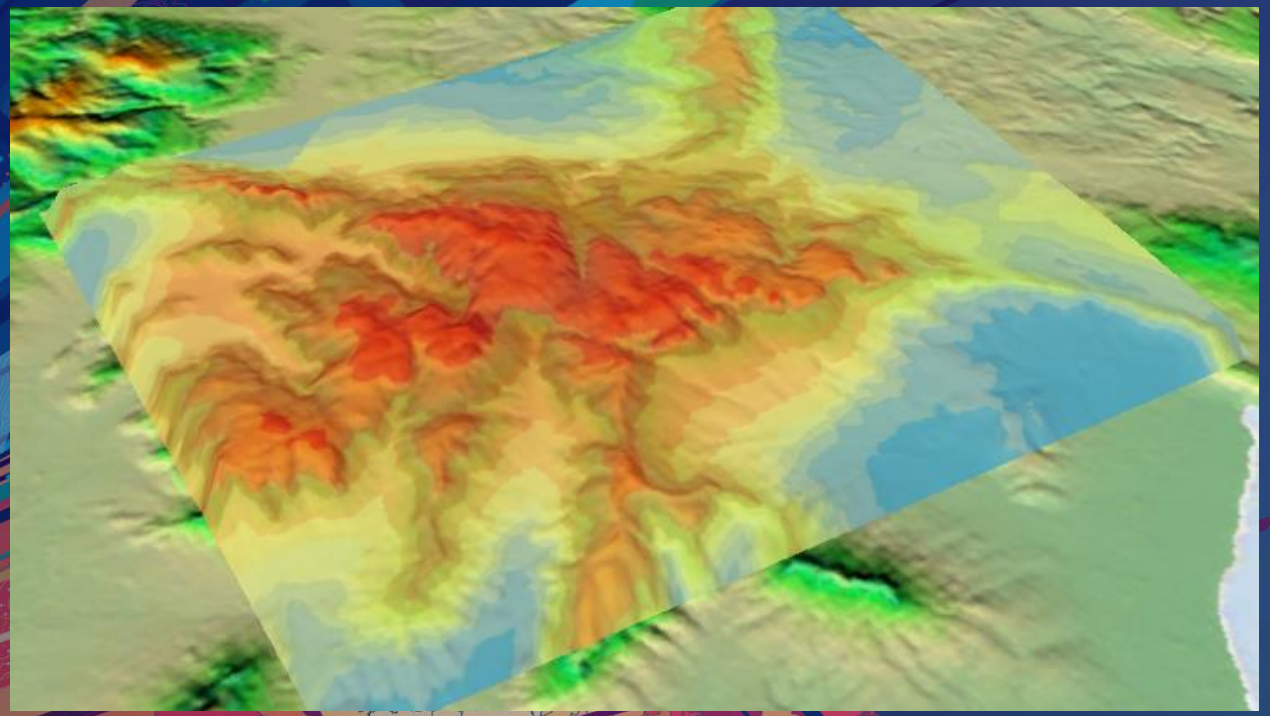
- Any realization might be the “real” thing!





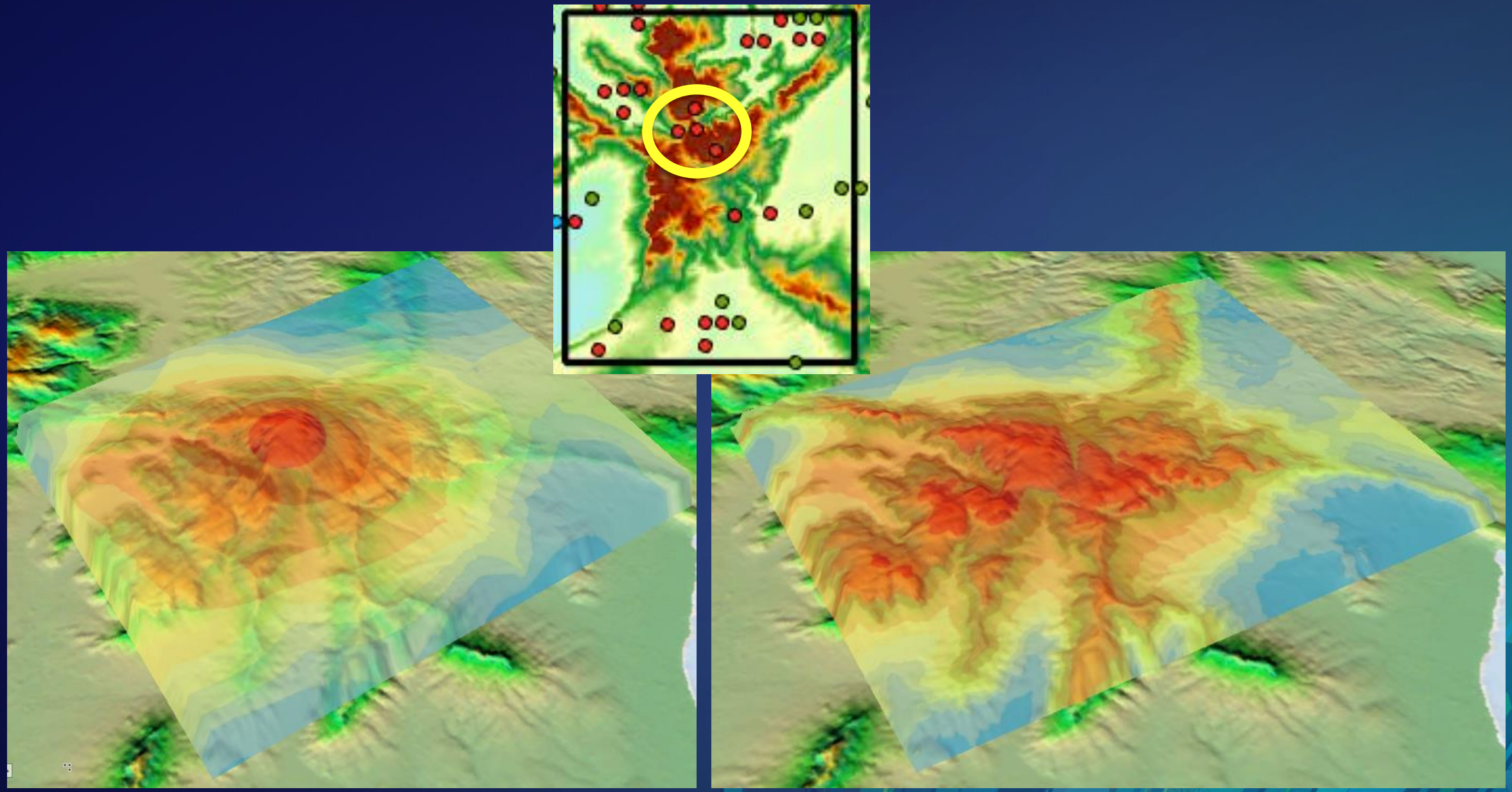






EBK Regression Prediction

Eric Krause



EBK Regression Prediction

Geostatistical layers

Eric Krause

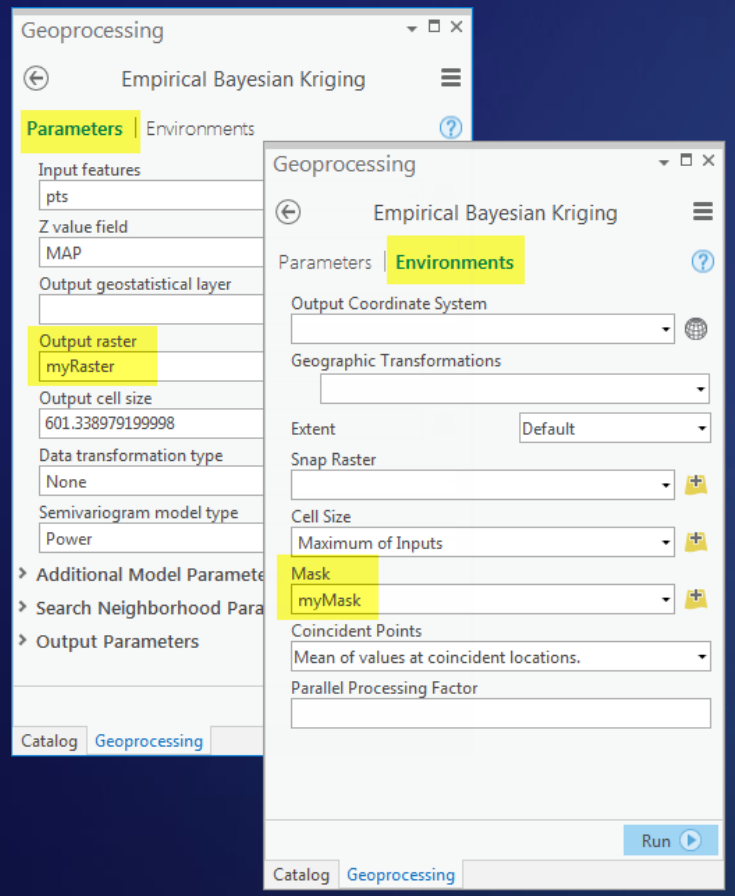
Where is GA used?

- Anyone who needs to statistically explore data and create surfaces for a number of variables will benefit from this statistical software package.

- Some of the various fields that use ArcGIS Geostatistical Analyst include:
  - agriculture,
  - geology,
  - meteorology,
  - hydrology,
  - archaeology,
  - forestry,
  - oceanography,
  - fishery,
  - health care, and
  - environmental studies.

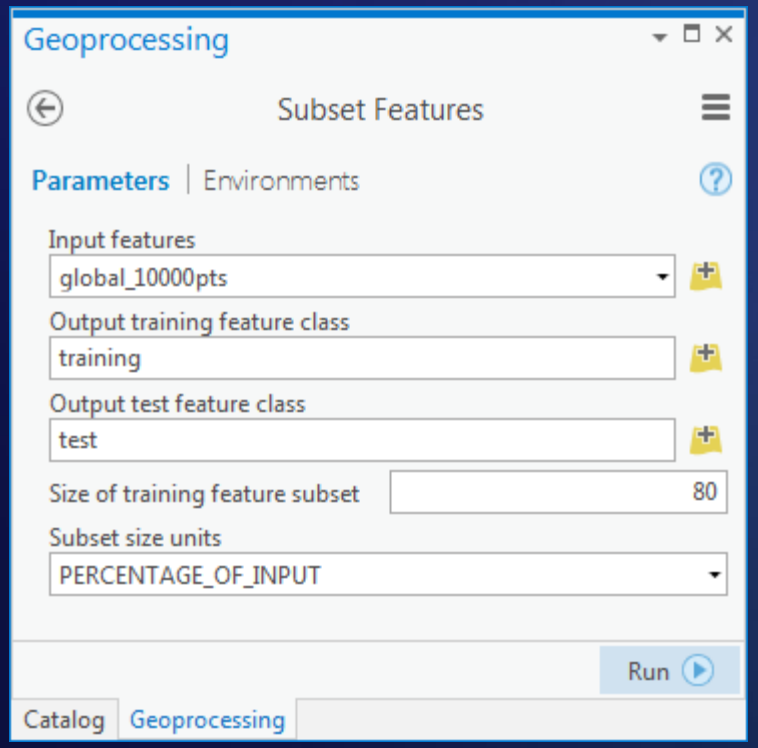
Tips & Tricks

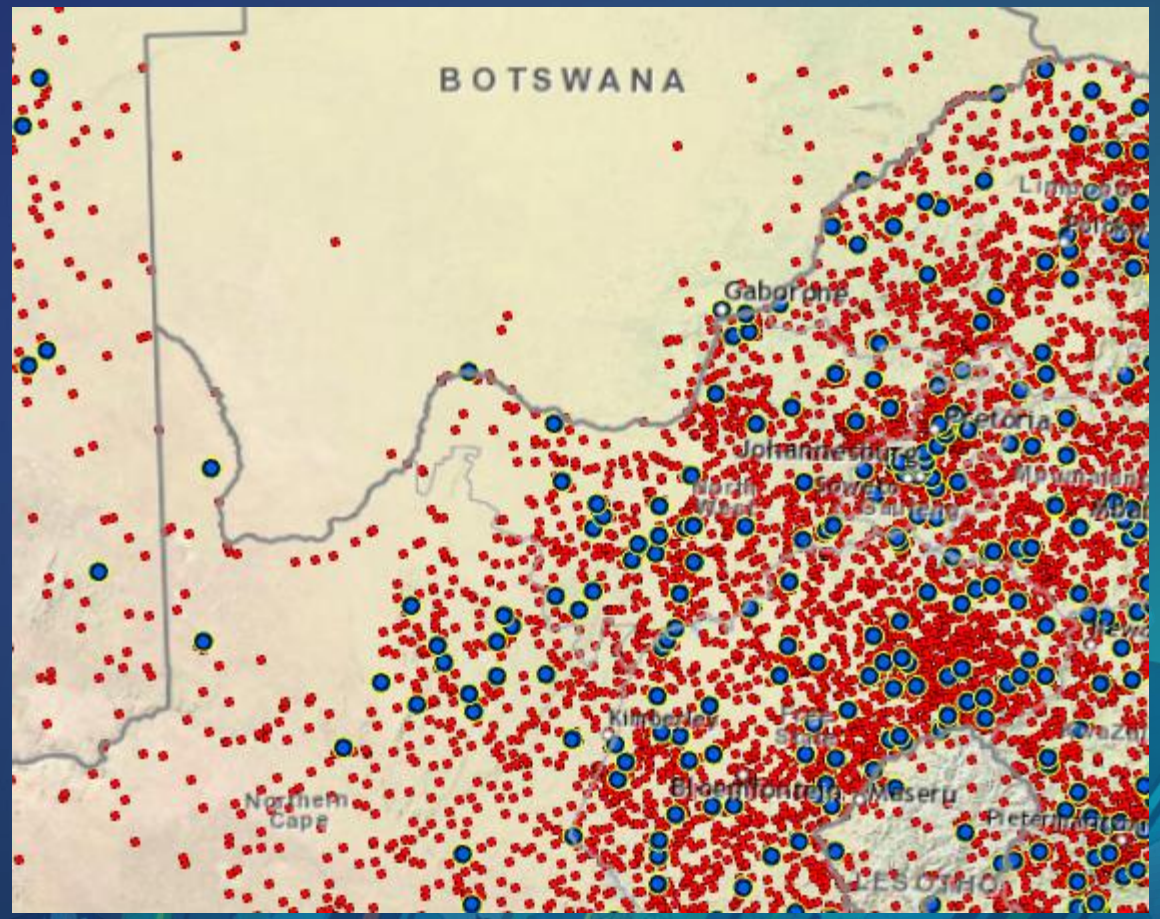
- Use Mask when creating a raster
  - 8700 pixels inside (55,000 outside) Japan (6 ½ times)



Tips & Tricks

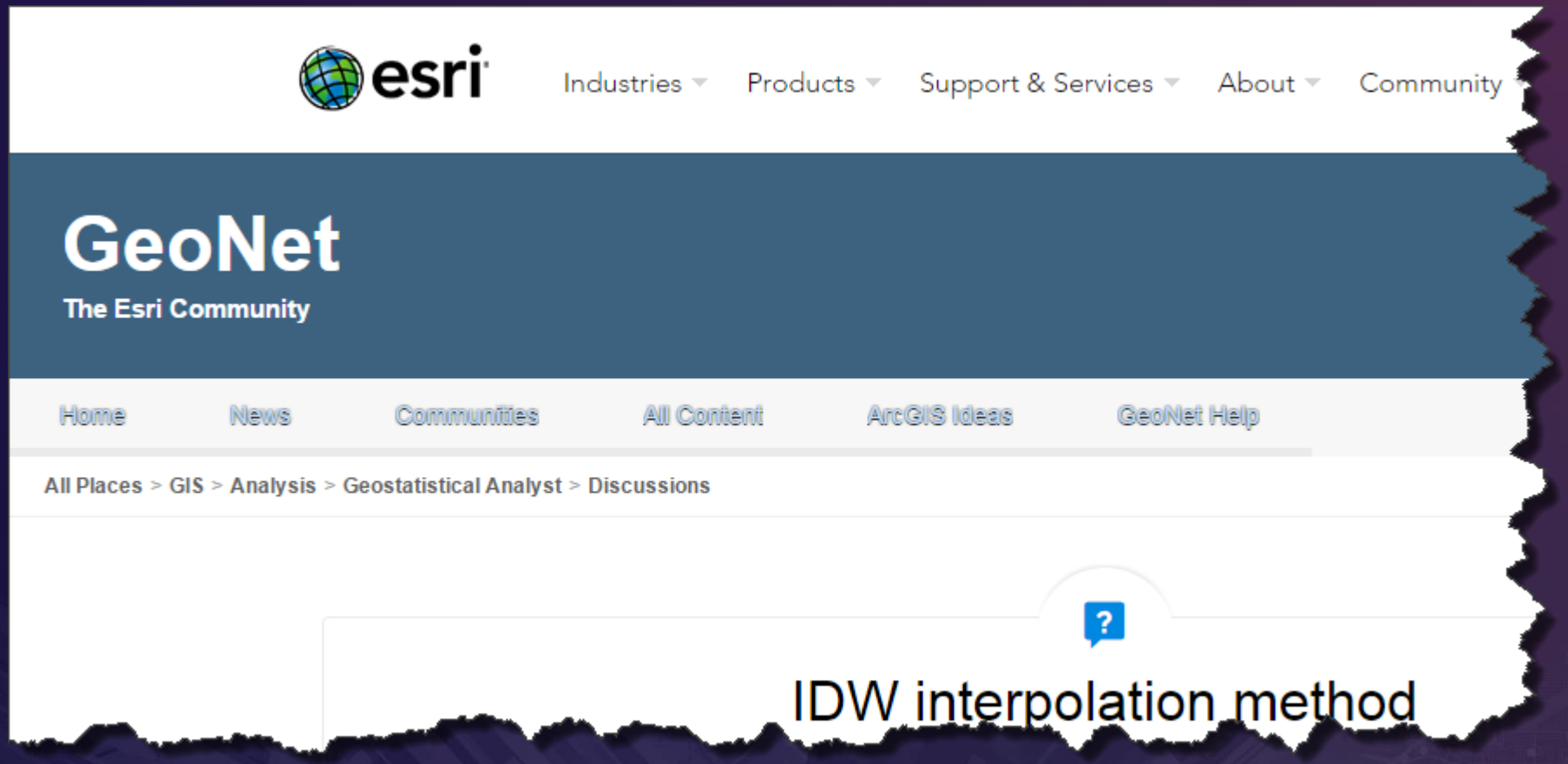
- Subset of the data
  - SubsetFeatures GP tool
  - Selection





Conclusions

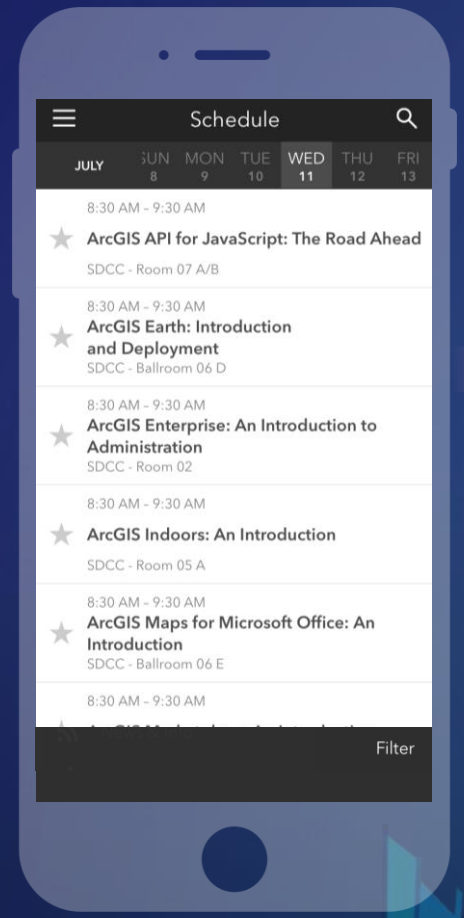
https://geonet.esri.com/



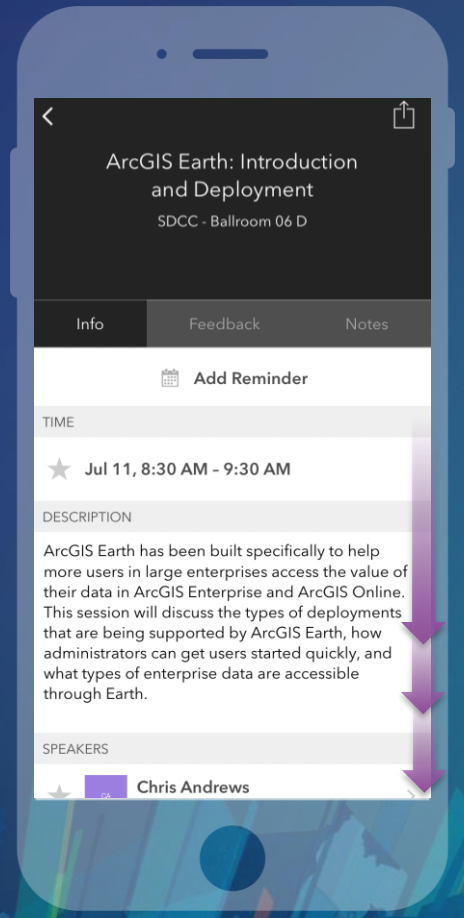
Please Take Our Survey on the App

Download the Esri Events app and find your event

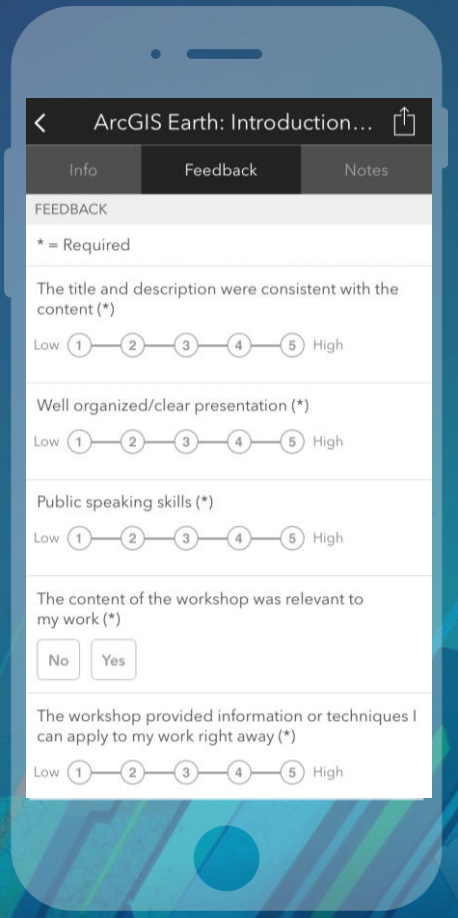
Select the session you attended



Scroll down to find the feedback section



Complete answers and select “Submit”



esri | THE SCIENCE OF WHERE

Find optimal script

```
iterationsList = [30, 100, 500, 1000]
subsetSizeList = [100, 200, 500]
overlapList = [1,3,5]
semivariogram_List =  [['POWER','NONE'], ['LINEAR','NONE'], ['THIN_PLATE_SPLINE','NONE'],
                       ['EXPONENTIAL_DETRENDED','EMPIRICAL'], ['EXPONENTIAL','EMPIRICAL'],
                       ['K_BESSEL_DETRENDED','EMPIRICAL'], ['K_BESSEL','EMPIRICAL'],
                       ['WHITTLE','EMPIRICAL'], ['WHITTLE_DETRENDED','EMPIRICAL']]

try:
    for iterations in iterationsList:
        for i in range(0, len(semivariogram_List)):
            for subsetSize in subsetSizeList:
                for overlap in overlapList:
                    svg = semivariogram_List[i][0]
                    transf=semivariogram_List[i][1]
                    t1 = time.time()
                    result = arcpy.EmpiricalBayesianKriging_ga(inFC, inField, outLyr, outRas, cellsize,
                                                               transf, subsetSize, overlap, iterations,
                                                               sn, 'PREDICTION', '', '', '', svg)
                    t2 = time.time()
                    tebk = t2-t1
                    cv = arcpy.CrossValidation_ga(outLyr)
                    rmseValue = cv.rootMeanSquare
                    rmsStd = cv.rootMeanSquareStandardized
                    toWriteList = [rmseValue, rmsStd, svg, transf, tebk, subsetSize, overlap, iterations]
                    outFile.writerow(toWriteList)
                    arcpy.Delete_management(outLyr)

except:
    print ('Tool execution FAILED')
    print (arcpy.GetMessages())
```