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

- **ArcGIS for Geostatistical Analyst: An Introduction (Tues 8:30-9:45 SDCC Rm17B)**
- Creating Surfaces from Various Data Sources (Tues 3:15-4:30 SDCC Rm09)
- Concepts and Applications of Kriging (Tues 3:15-4:30 SDCC Rm17B)
- *Empirical Bayesian Kriging and EBK Regression Prediction – Robust Kriging as GP Tools (Tues 5:30-6:15 Th07)*

Wednesday

- *Choosing the Best Kriging Model for Your Data (Wed 11:30-12:15 SDCC Th07)*
- **ArcGIS for Geostatistical Analyst: An Introduction (Wed 1:30-2:45 SDCC Rm17B)**

Thursday

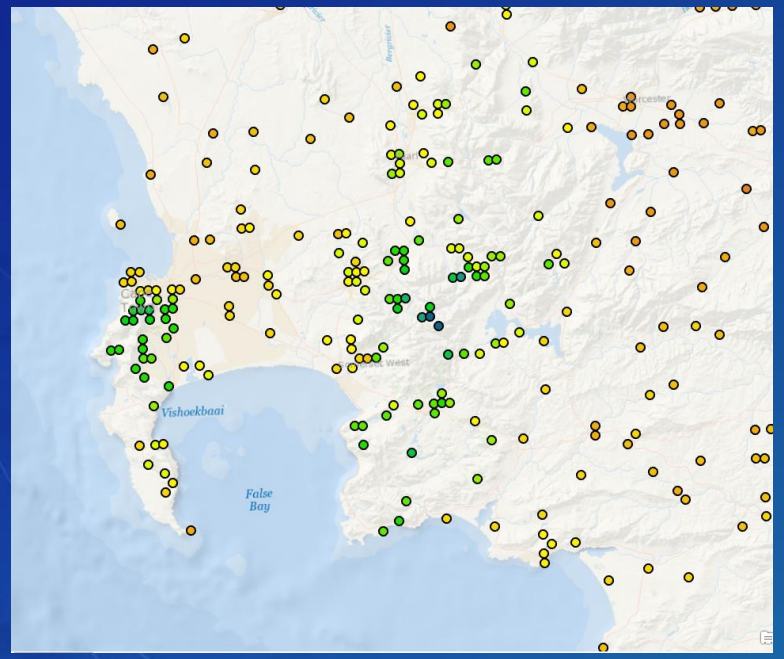
- Geostatistics in Practice: Learning Kriging Through Examples (Thurs 8:30-9:45 SDCC Rm10)
- *Surface Interpolation in ArcGIS (Thurs 10:30-11:15 SDCC Th07)*
- *Performing Polygon-to-Polygon Predictions using Areal Interpolation (Thurs 11:30-12:15 Th07)*
- Creating Surfaces from Various Data Sources (Thurs 3:15-4:30 SDCC Rm09)

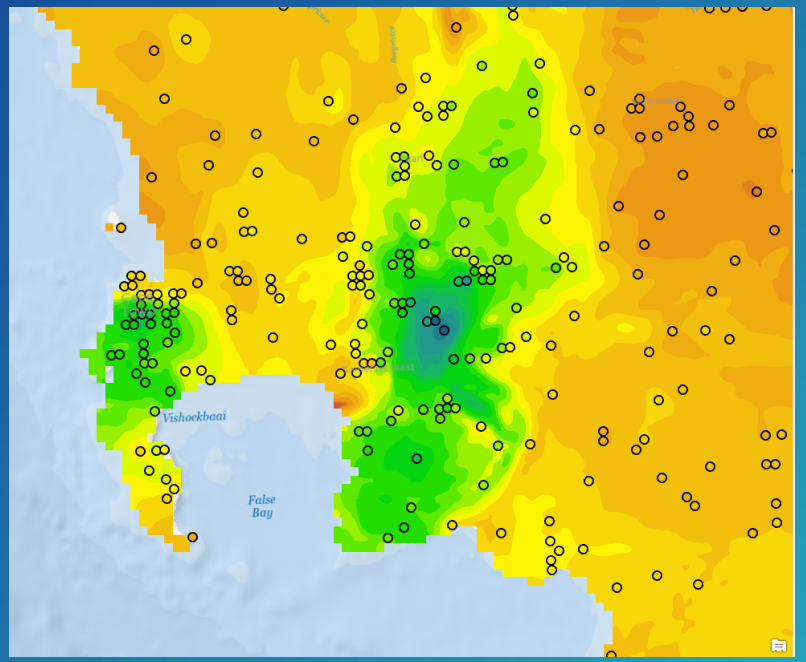
What is geostatistics?

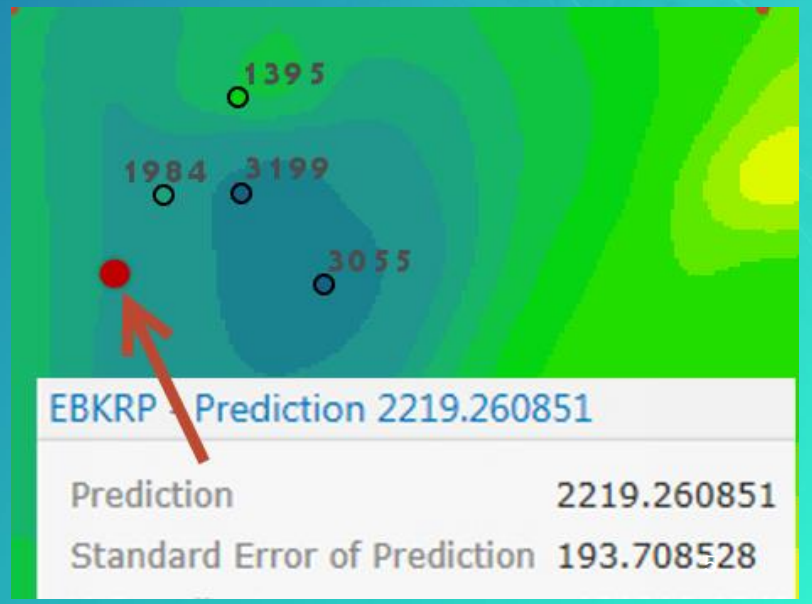
- is a class of statistics used to analyze and predict values associated with spatial phenomena.
- it incorporates the spatial coordinates of the data
- Has evolved to not only provide
  - interpolated values, but also
  - measures of uncertainty

ArcGIS for Geostatistical Analyst

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

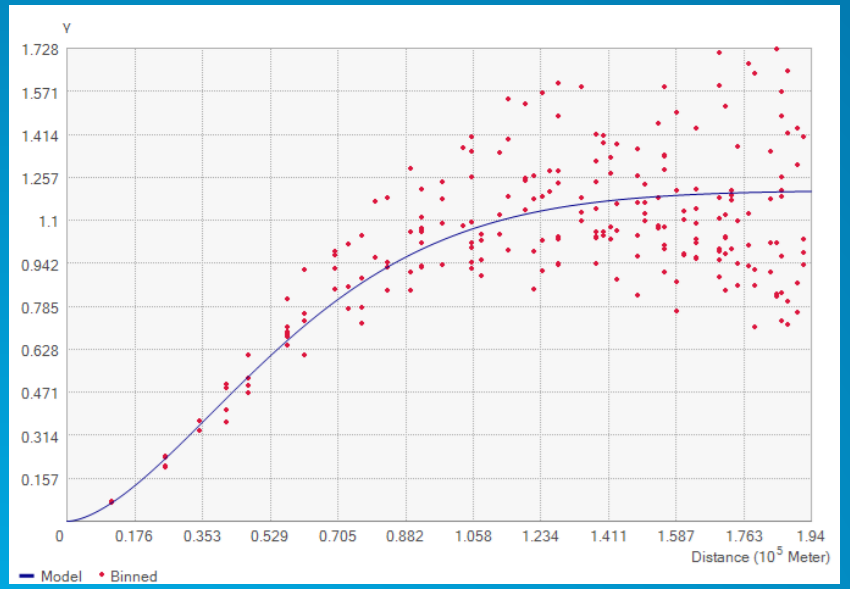


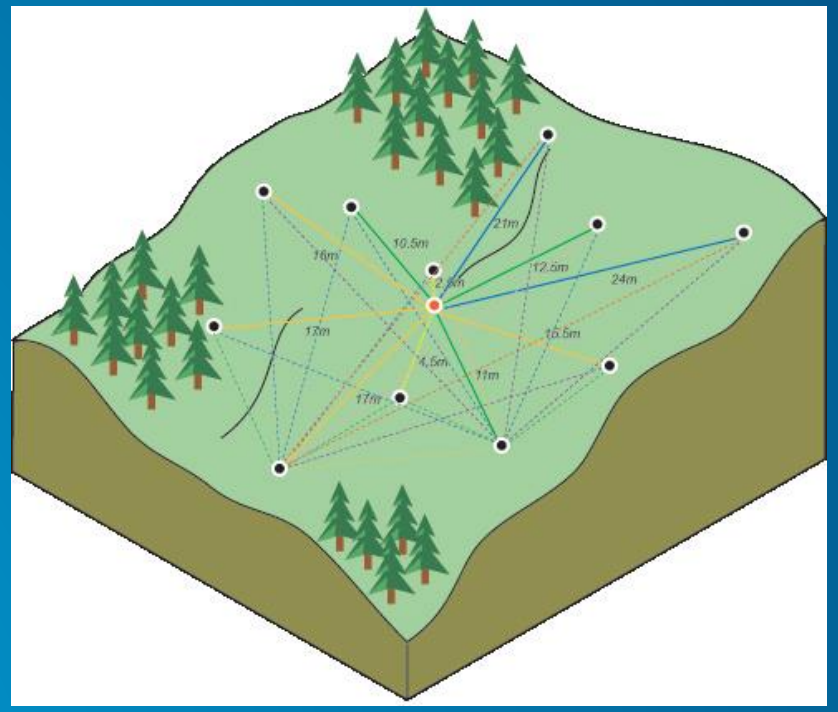


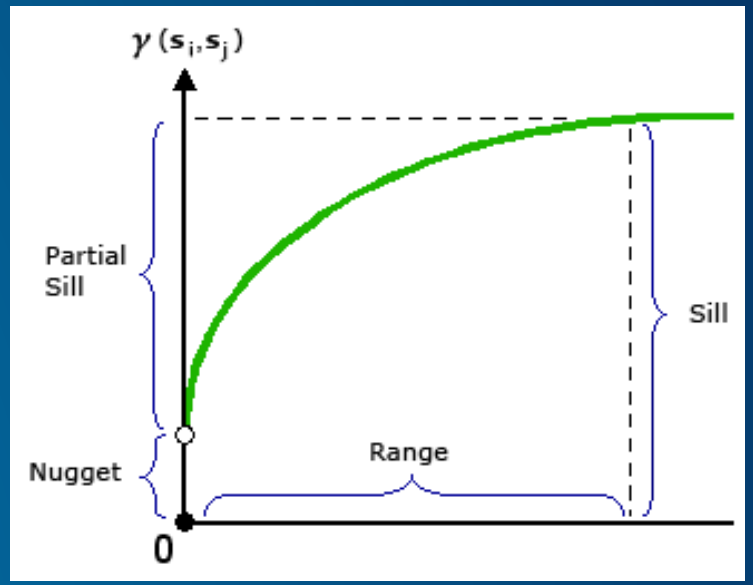


What is a semivariogram?

Semivariogram(distance h) = 0.5 * average [$(value_i - value_j)^2$]

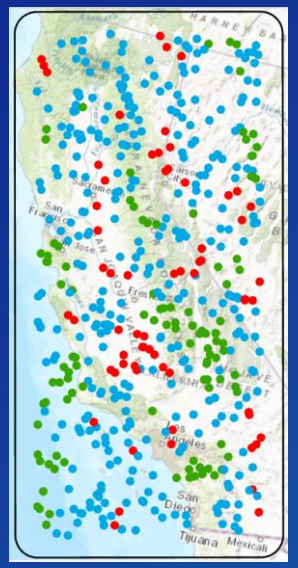


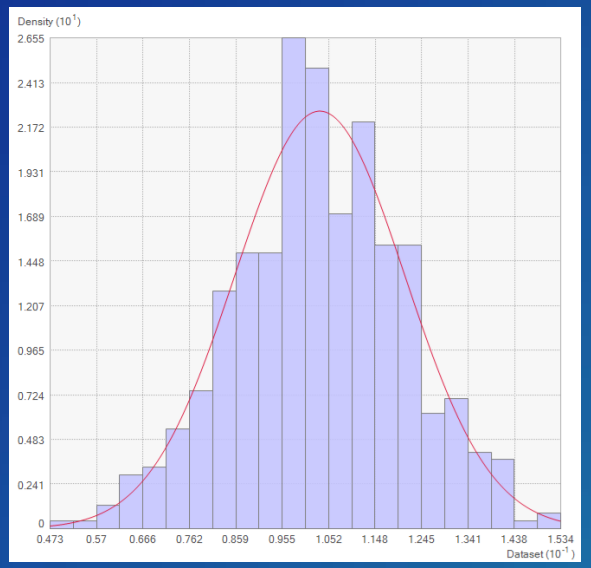


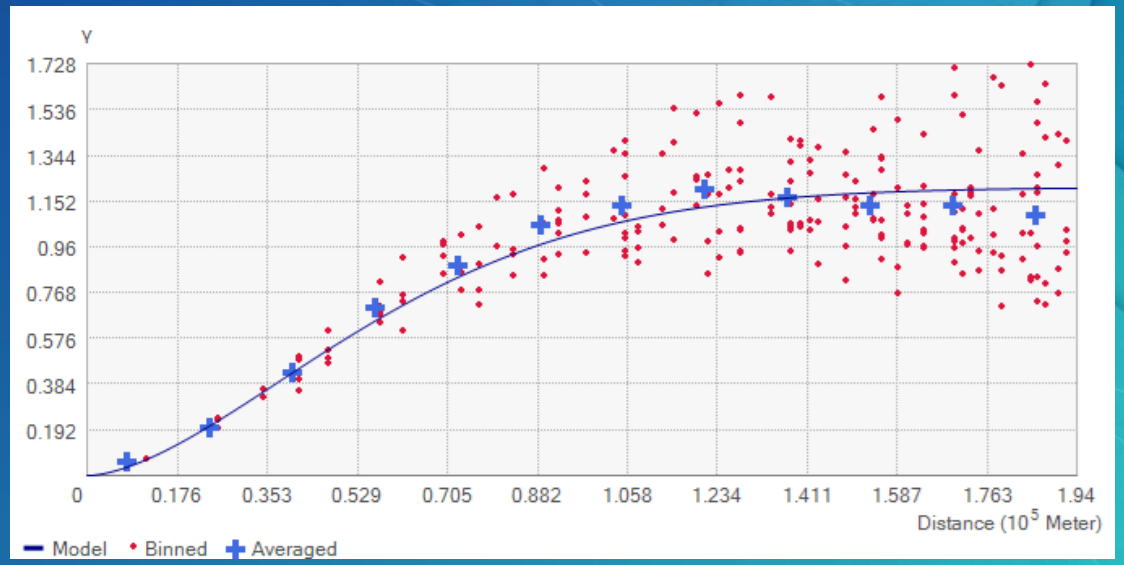


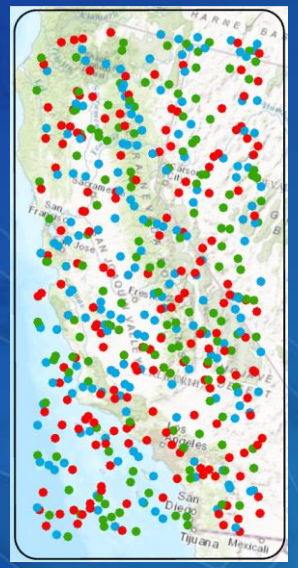
- Range = separation distance between pairs
- Sill = plateau the variogram reaches at the range
- Nugget = sampling error and short scale variability

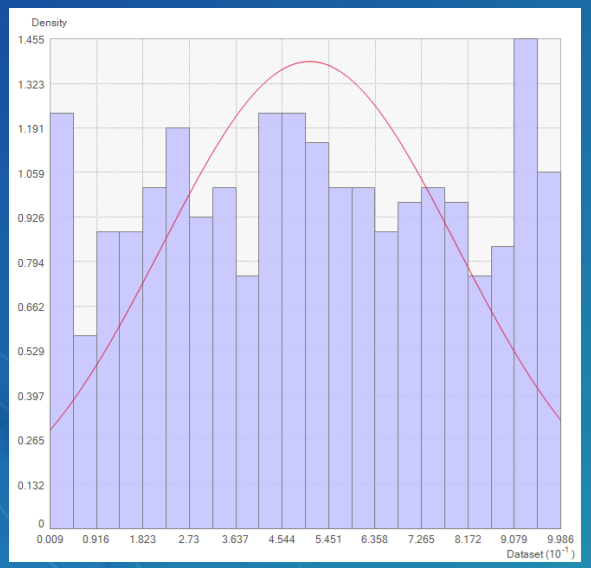
Spatial autocorrelation

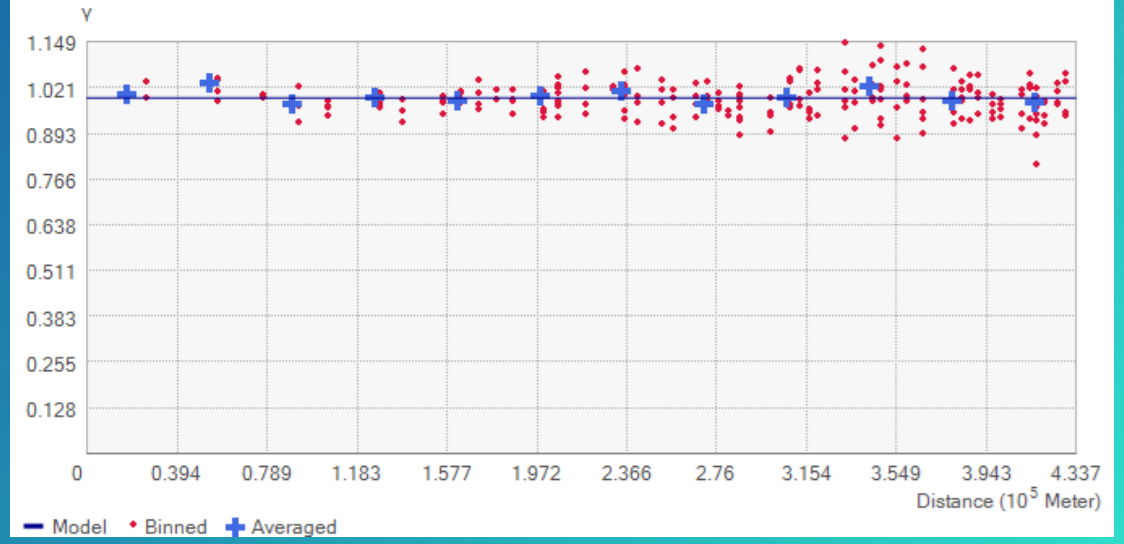








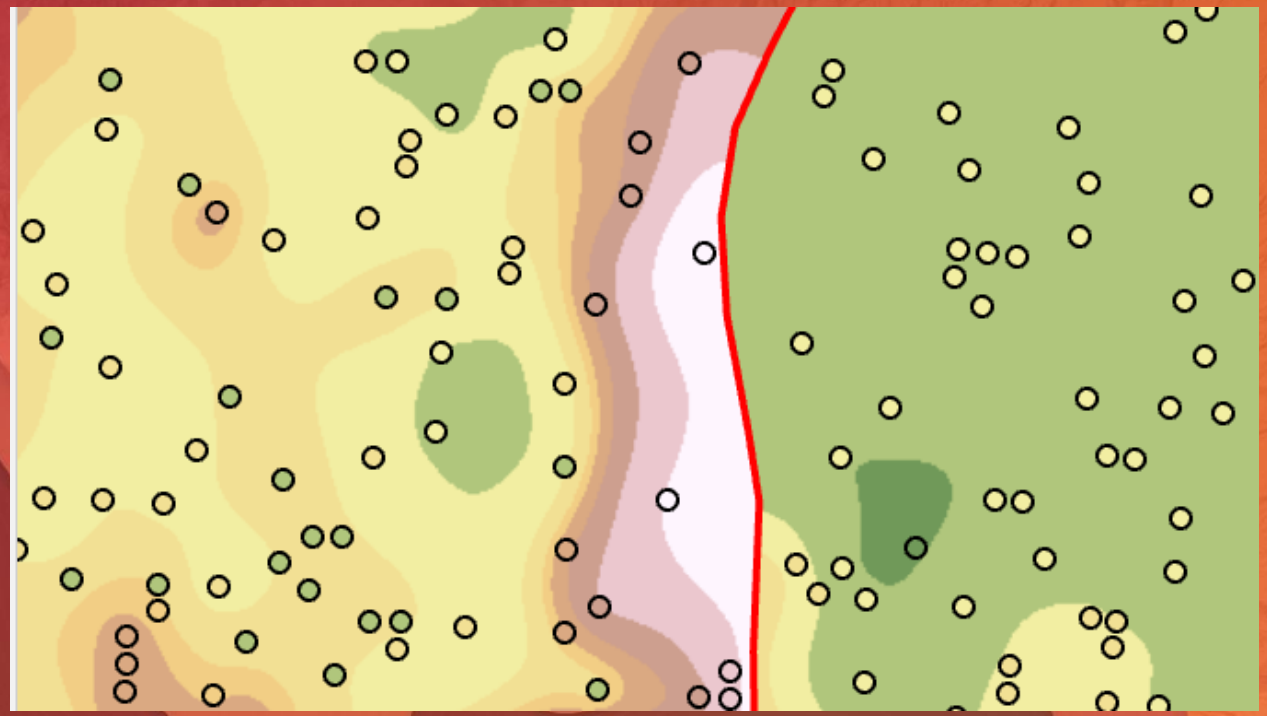




Geostatistical Analyst – What is it?

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

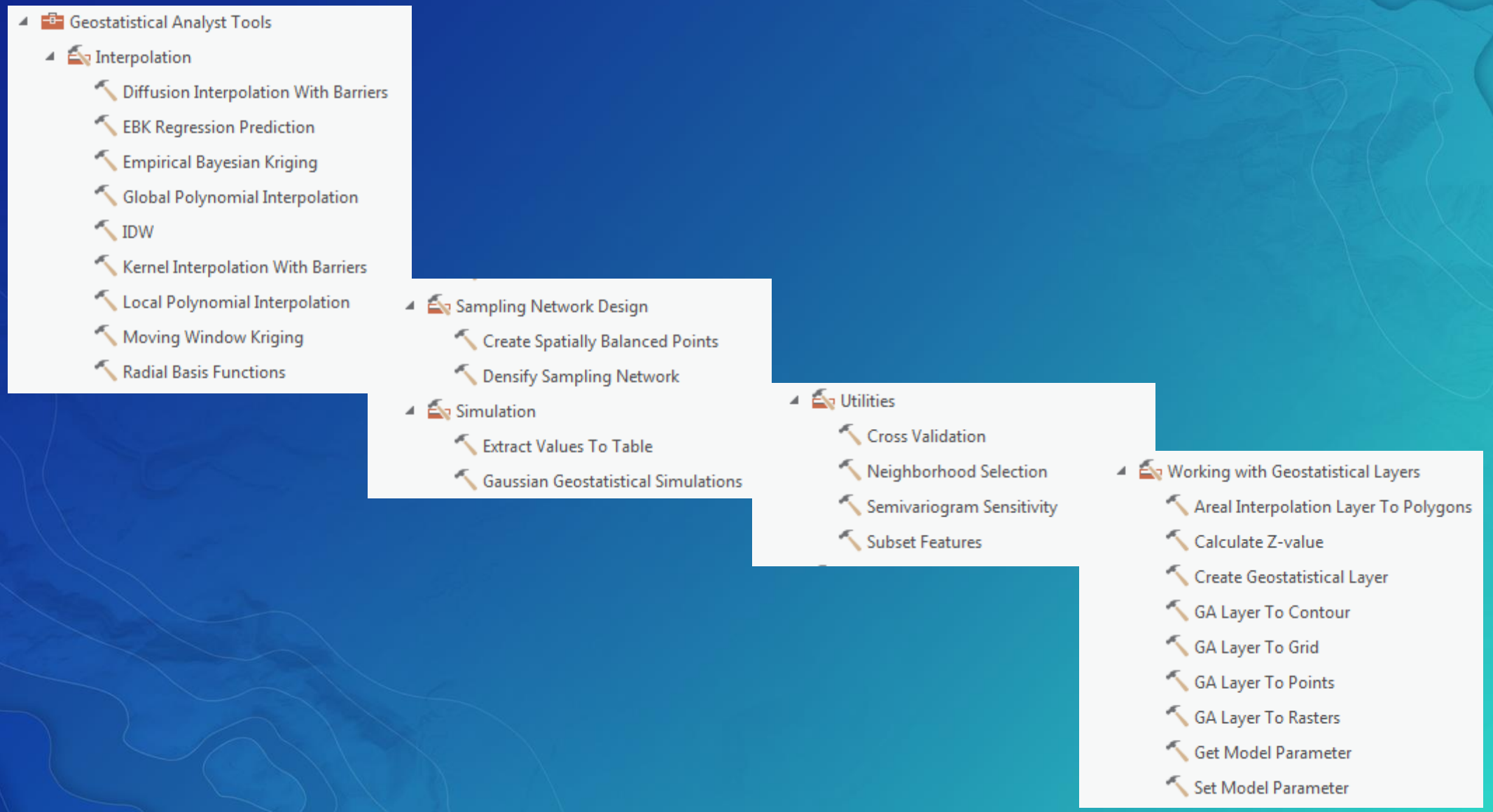


GP tool

Kernel Interpolation with Barriers

Eric Krause

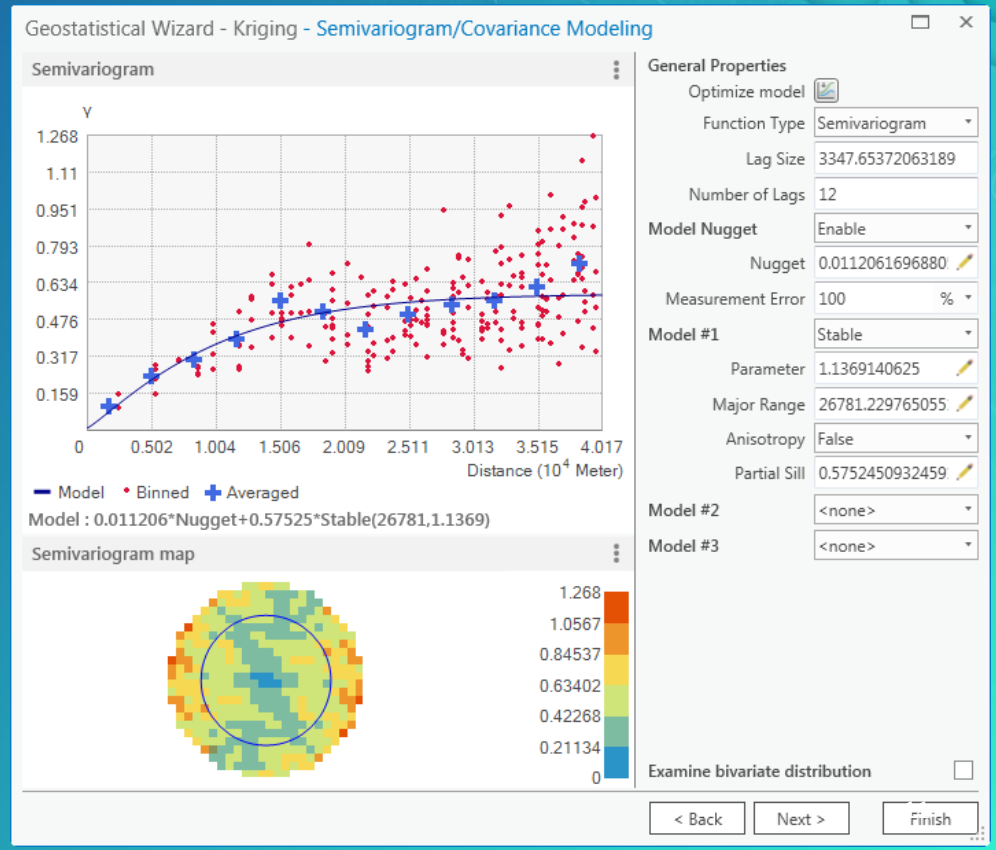
Geostatistical Analyst – Geoprocessing tools

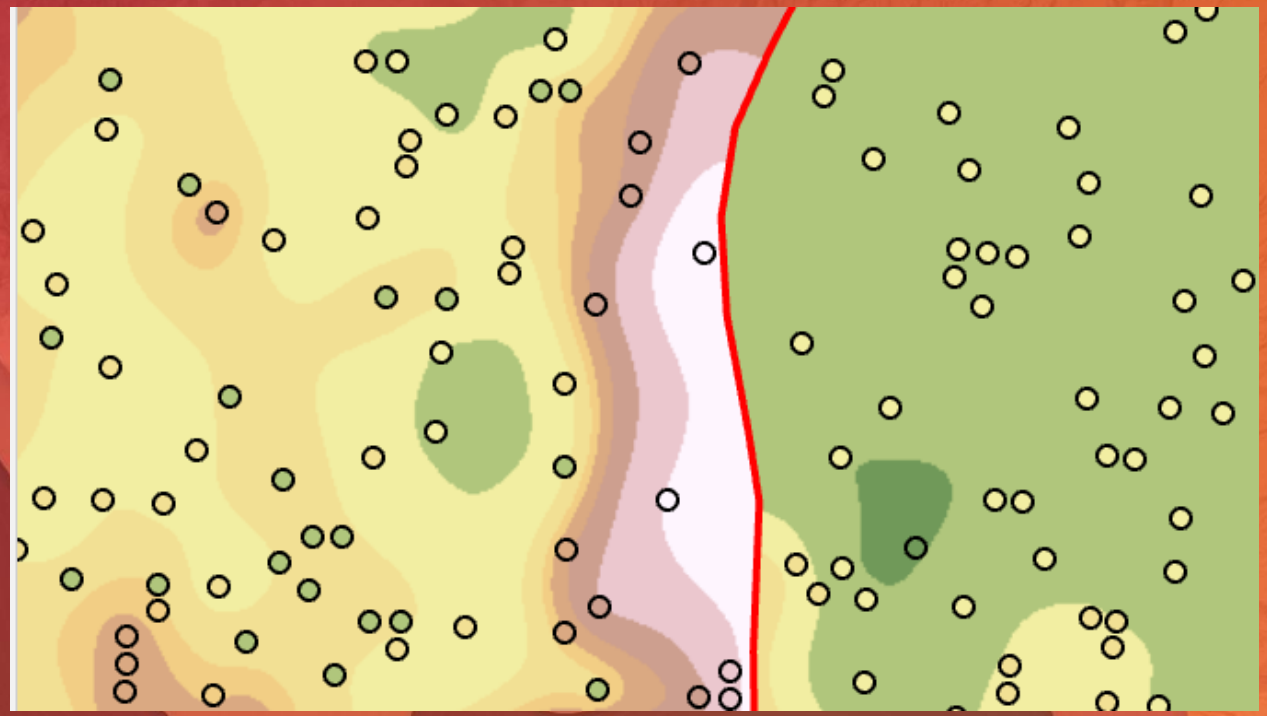


Geostatistical Analyst – What is it?

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

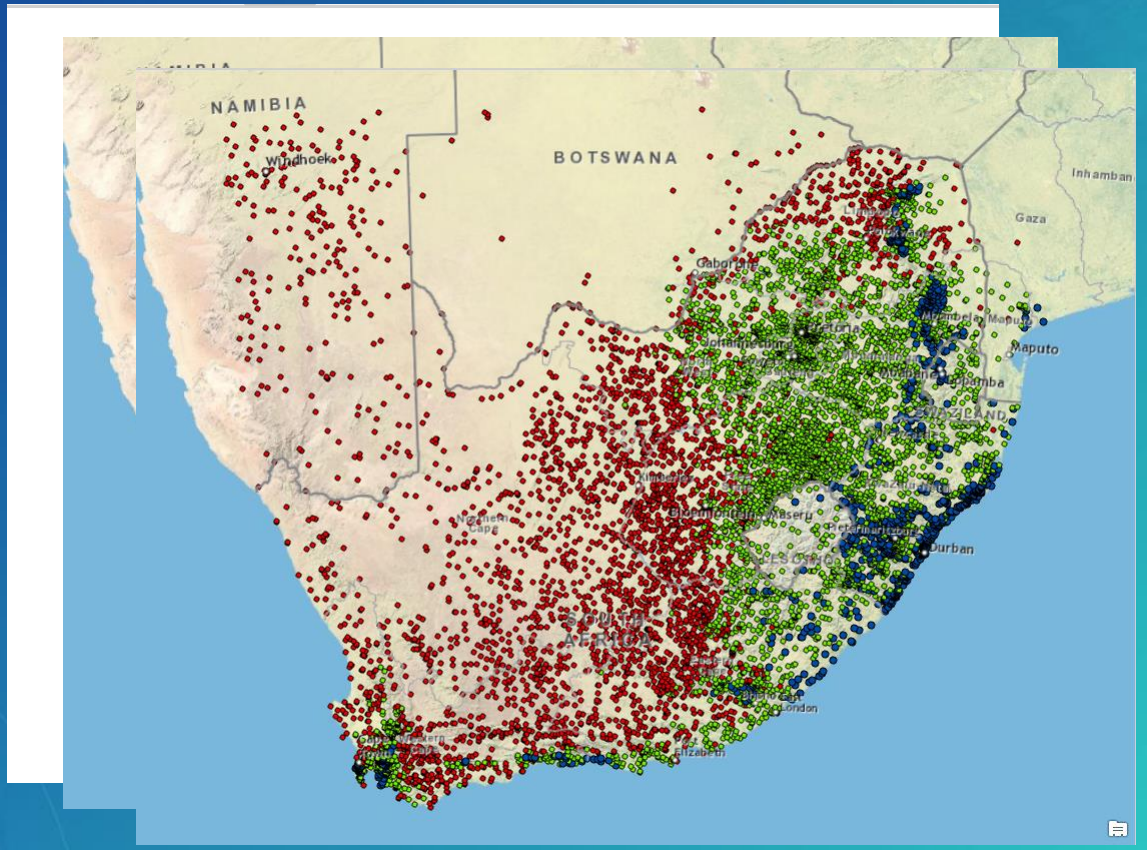
Kernel Interpolation with Barriers

Eric Krause

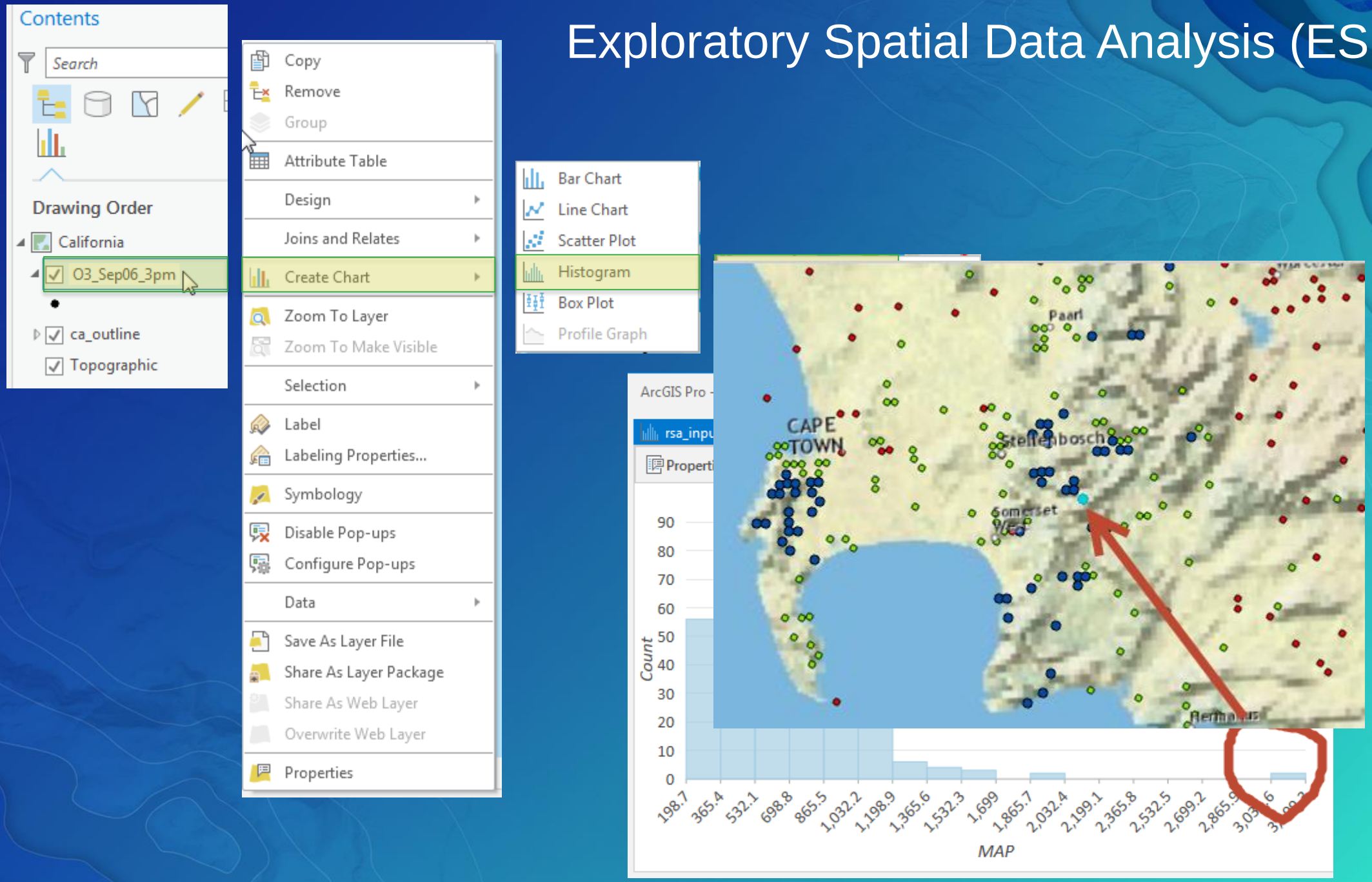
ESDA

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?

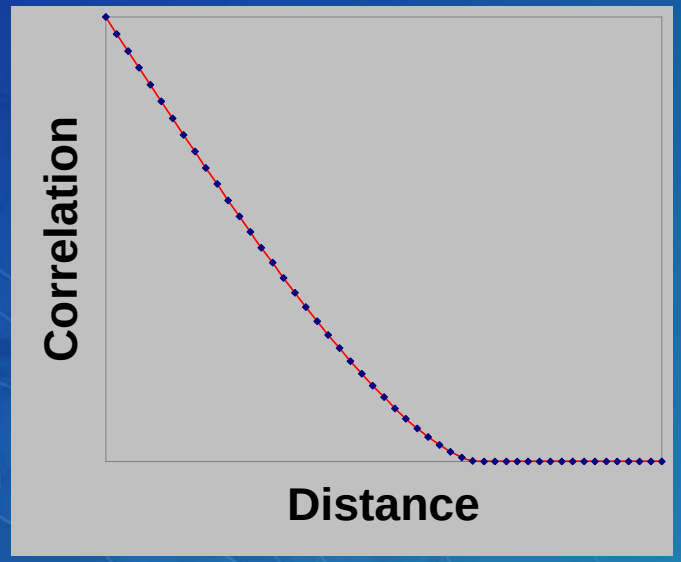


Exploratory Spatial Data Analysis (ESDA)



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



EBK Regression Prediction

Empirical Bayesian Kriging

Geostatistical Wizard - Kriging

Ordinary Kriging
- Prediction
- Quantile
- Probability
- Prediction Standard Error

Simple Kriging
- **Prediction**
- Quantile
- Probability
- Prediction Standard Error

Universal Kriging
- Prediction
- Quantile
- Probability
- Prediction Standard Error

Indicator Kriging
- Probability
- Standard Error of Indicators

Probability Kriging
- Probability
- Standard Error of Indicators

Disjunctive Kriging
- Prediction
- Probability
- Prediction Standard Error
- Standard Error of Indicators

What is kriging?

Tuesday

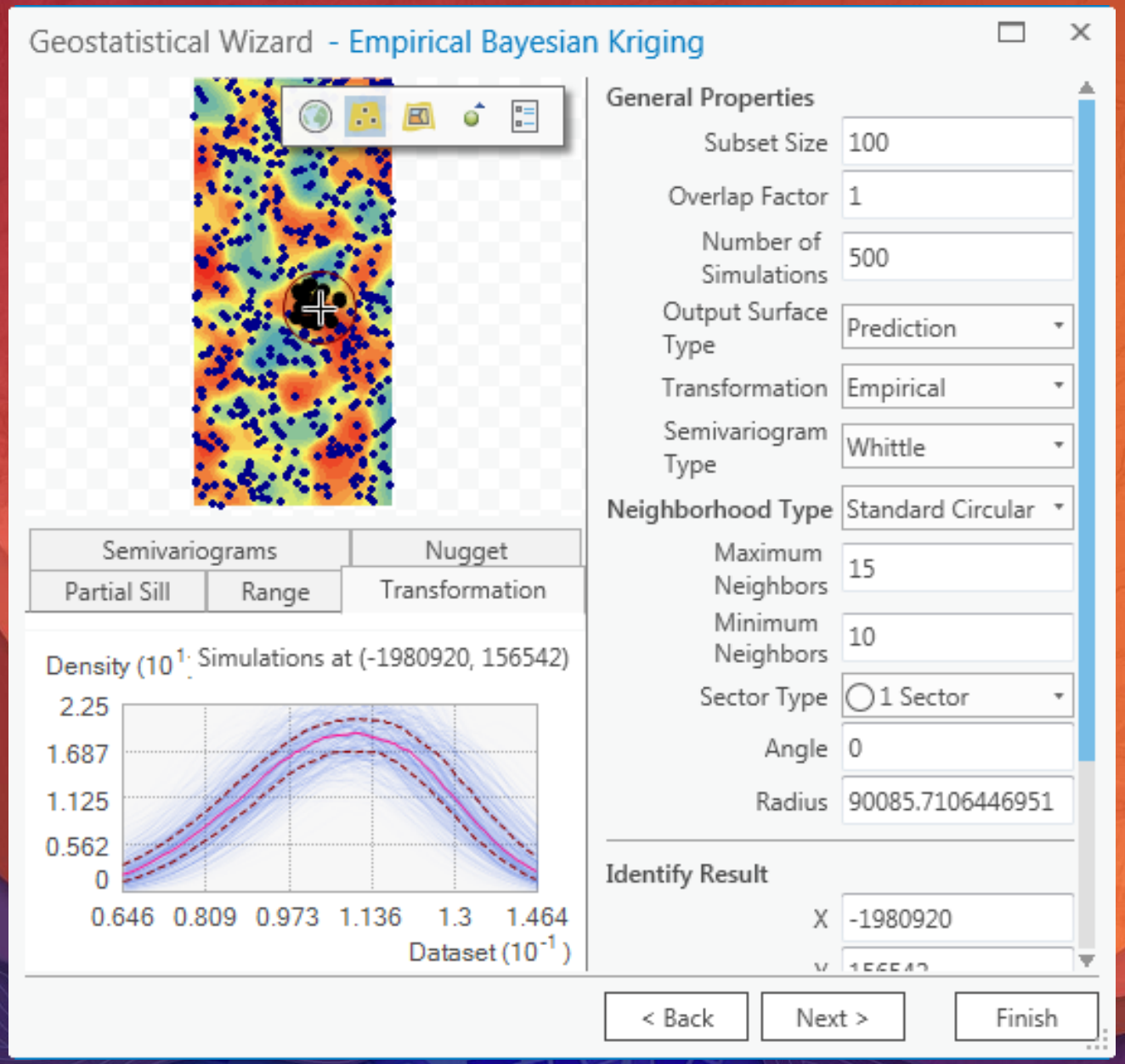
- Concepts and Applications of Kriging (Tues 3:15-4:30 SDCC Rm17B)
- *EBK and EBK Regression Prediction – Robust Kriging as GP Tools (Tues 5:30-6:15 Th07)*

Wednesday

- *Choosing the Best Kriging Model for Your Data (Wed 11:30-12:15 SDCC Th07)*

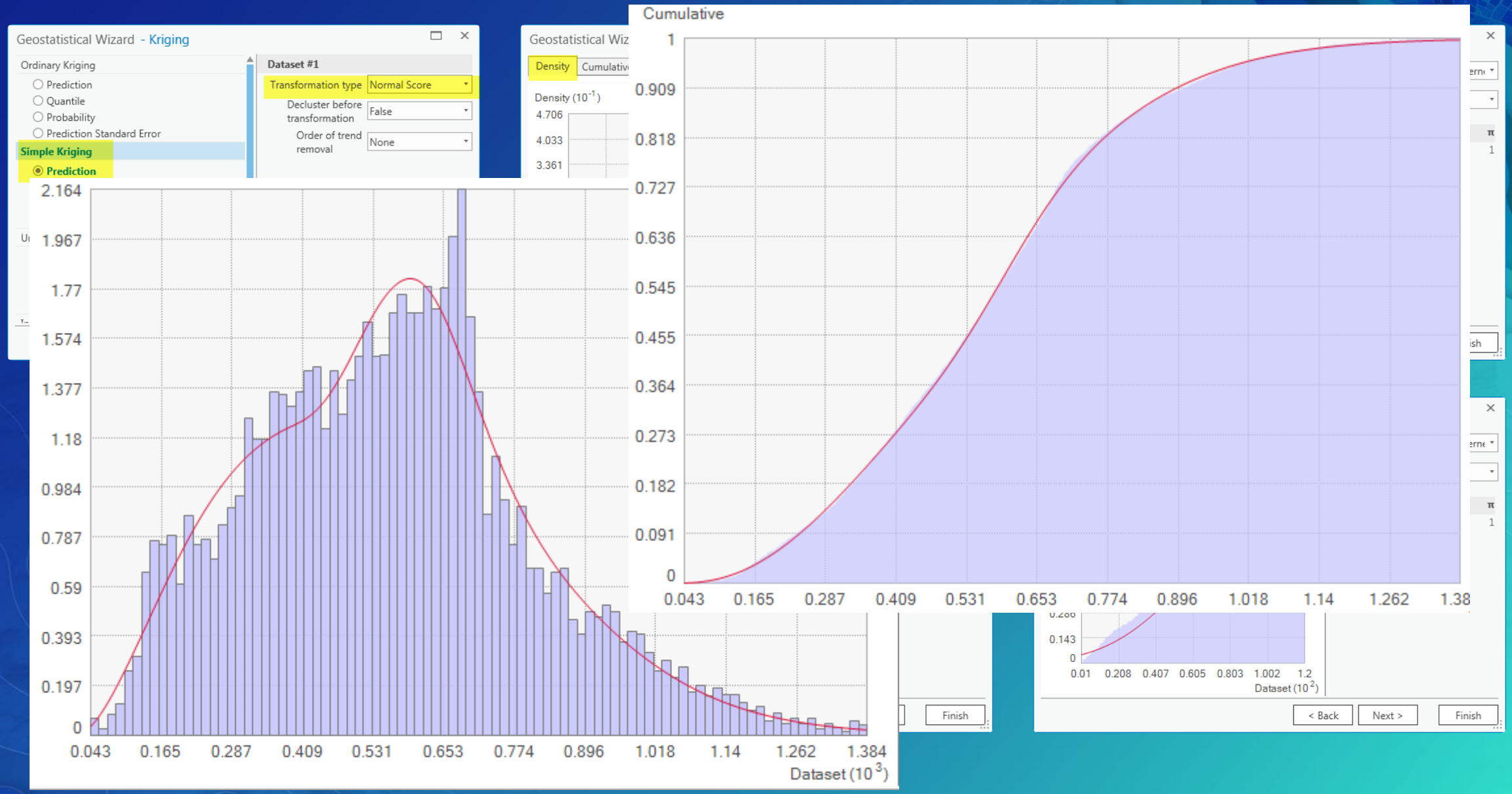
Thursday

- Geostatistics in Practice: Learning Kriging Through Examples (Thurs 8:30-9:45 SDCC Rm10)
- *Performing Polygon-to-Polygon Predictions using Areal Interpolation (Thurs 11:30-12:15 Th07)*



Geostatistical Wizard

Eric Krause


Geostatistical Wizard - Kriging
Ordinary Kriging
Prediction
Quantile
Probability
Prediction Standard Error
Simple Kriging
Prediction
Dataset #1
Transformation type
Normal Score
Decluster before transformation
False
Order of trend removal
None
Density
Cumulative
Density (10^{-1})
4.706
4.033
3.361
2.164
1.967
1.77
1.574
1.377
1.18
0.984
0.787
0.59
0.393
0.197
0
0.043 0.165 0.287 0.409 0.531 0.653 0.774 0.896 1.018 1.14 1.262 1.384
Dataset (10^{3})
Cumulative
1
0.909
0.818
0.727
0.636
0.545
0.455
0.364
0.273
0.182
0.091
0
0.043 0.165 0.287 0.409 0.531 0.653 0.774 0.896 1.018 1.14 1.262 1.38
0.143
0
0.01 0.208 0.407 0.605 0.803 1.002 1.2
Dataset (10^{2})
Finish
< Back
Next >
Finish

More ESDA

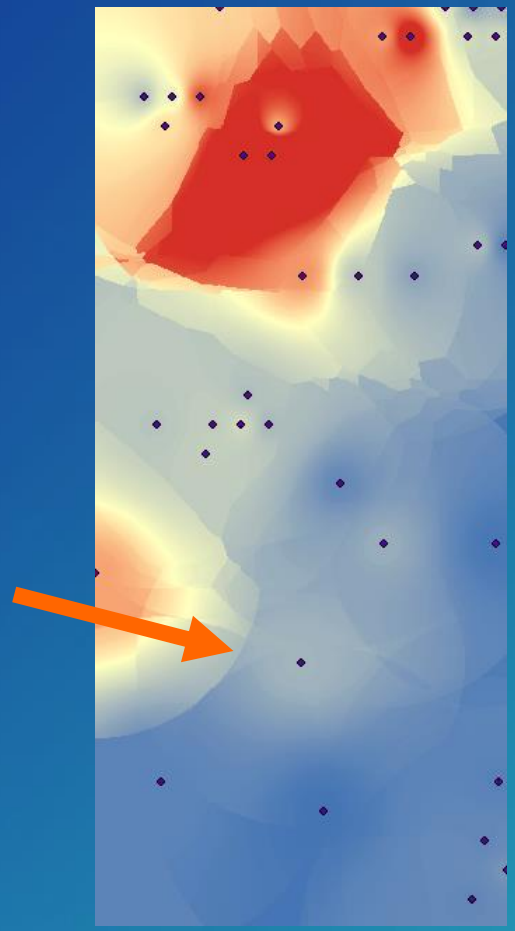
Interpolation workflow

- ESDA
- Interpolate
- Goodness of fit

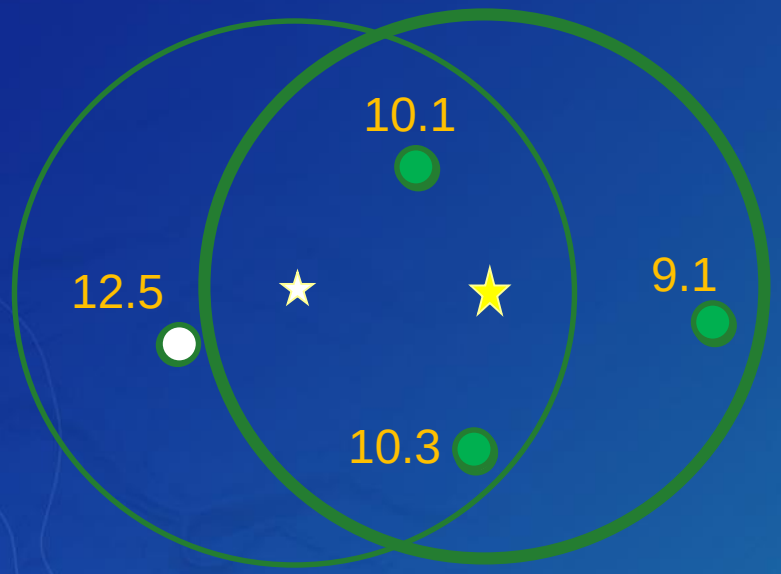
Why use ESRI's Geostatistical Analyst?

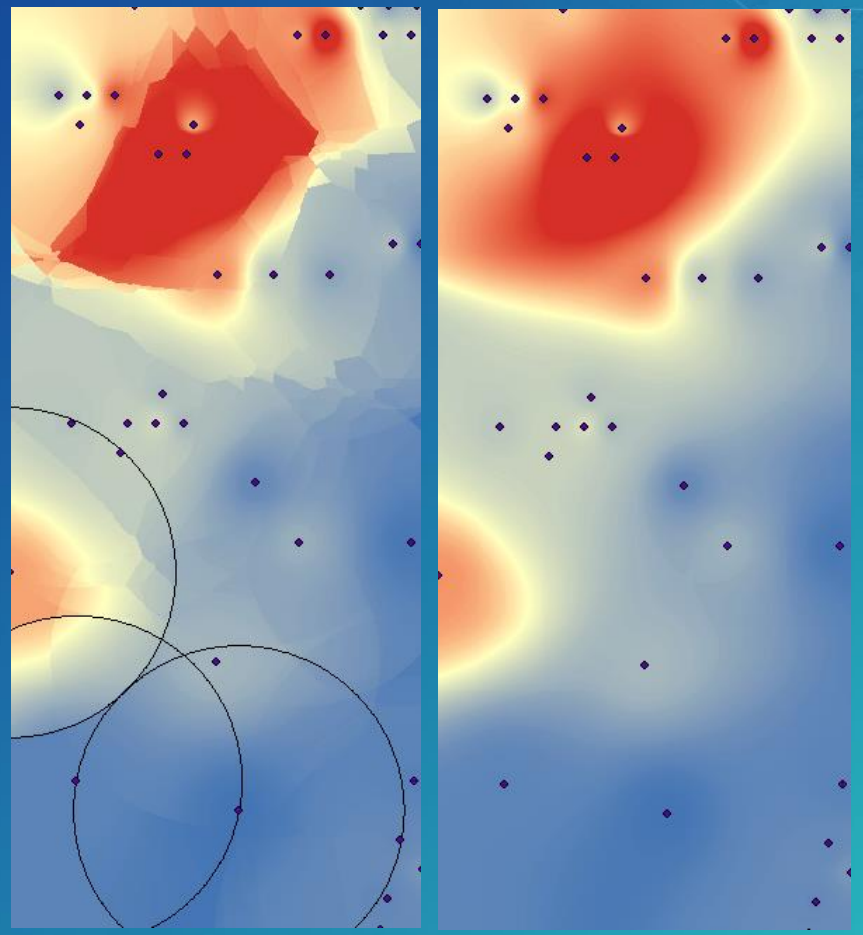
- Search neighborhood
  - Sectors
  - Smooth
- Chordal distance
- Cross validation
- Error maps
- Interactive Variography
- Barriers
- Simulations

Search neighborhood - Smooth



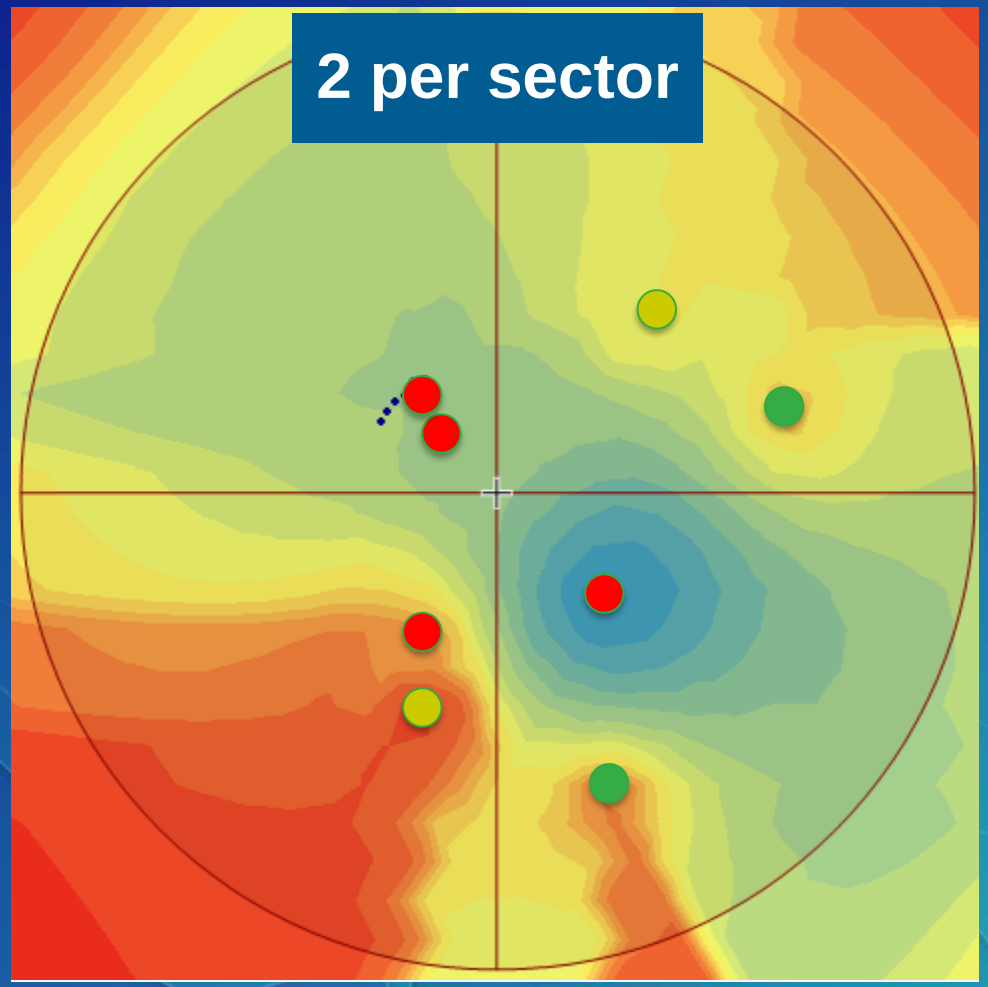
Search neighborhood - Smooth

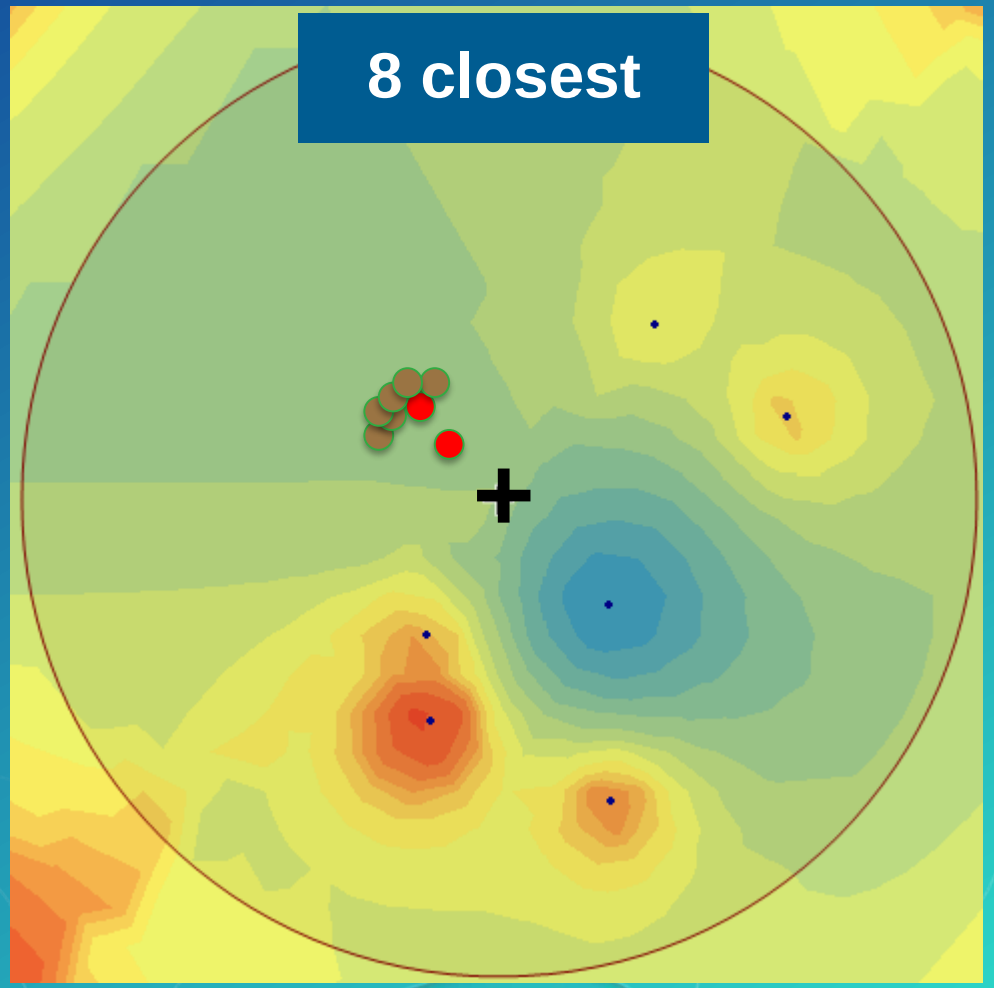




Unlike smoothing the output,
this method modifies the weights

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.

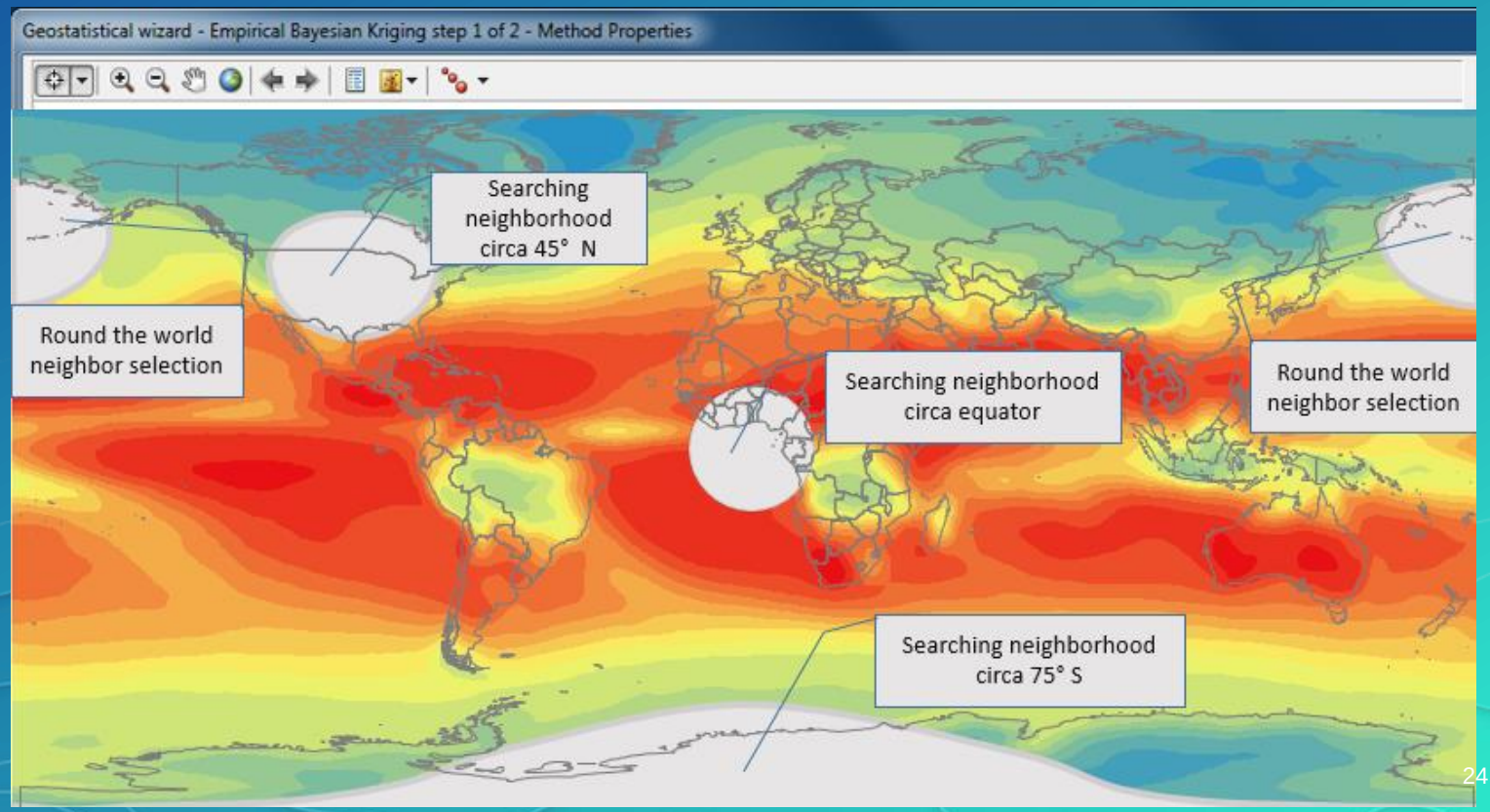
Distance between LA and New York

Geodesic = 3,939.1 km

Chordal = 3,877.0 km

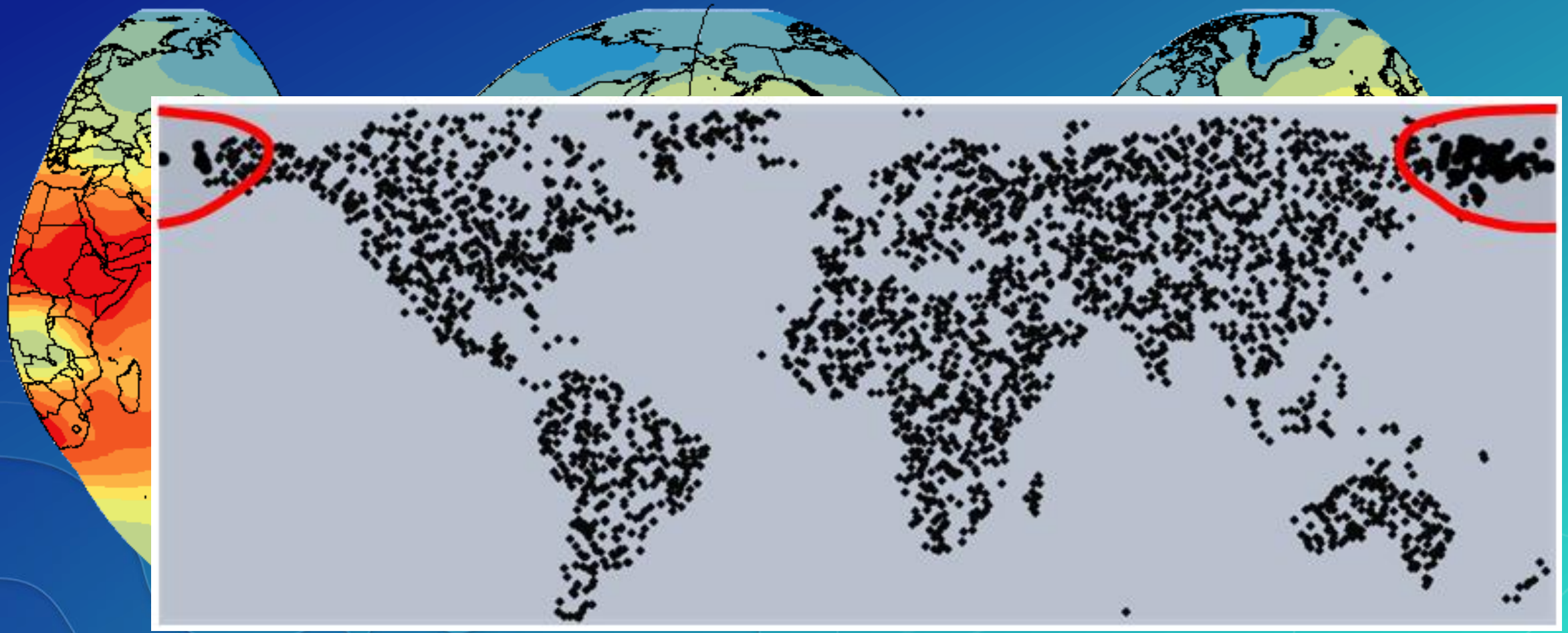
Difference = 62.1 km (1.5%)

Speed!

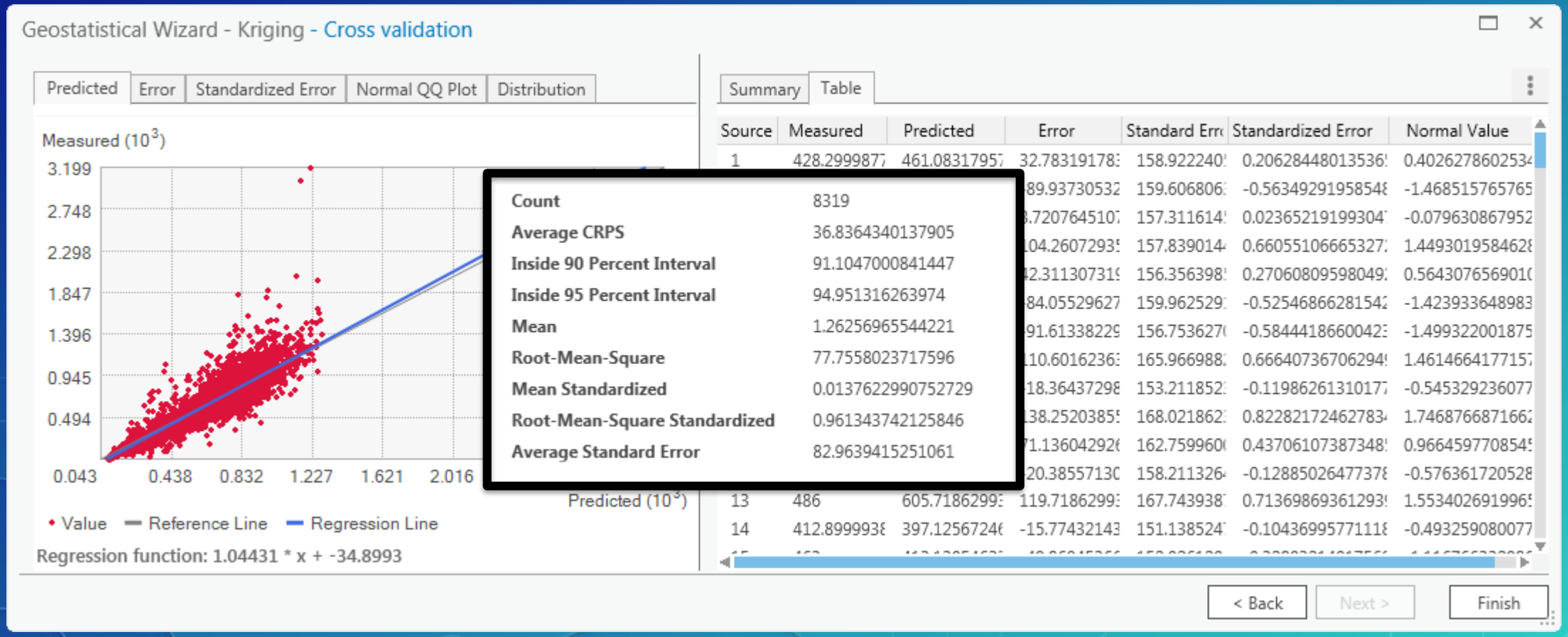


Chordal distances

Only for EBK and EBK Regression Prediction

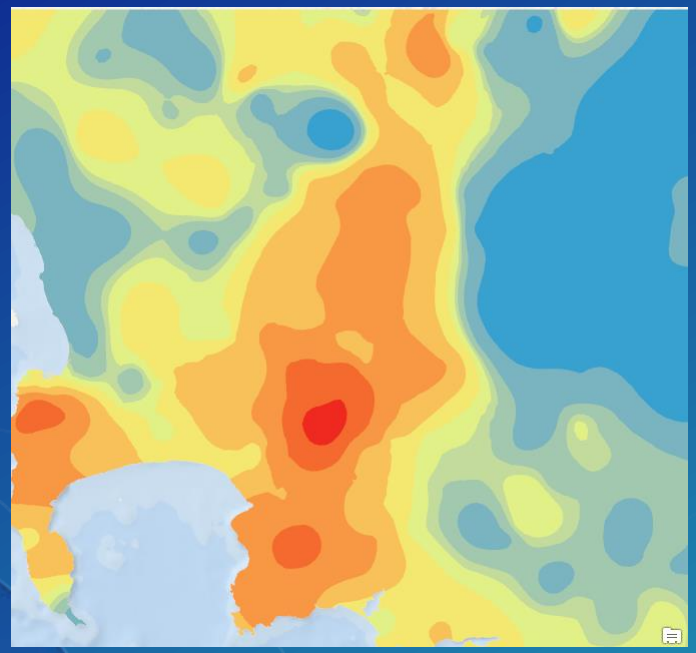


Cross validation / Validation

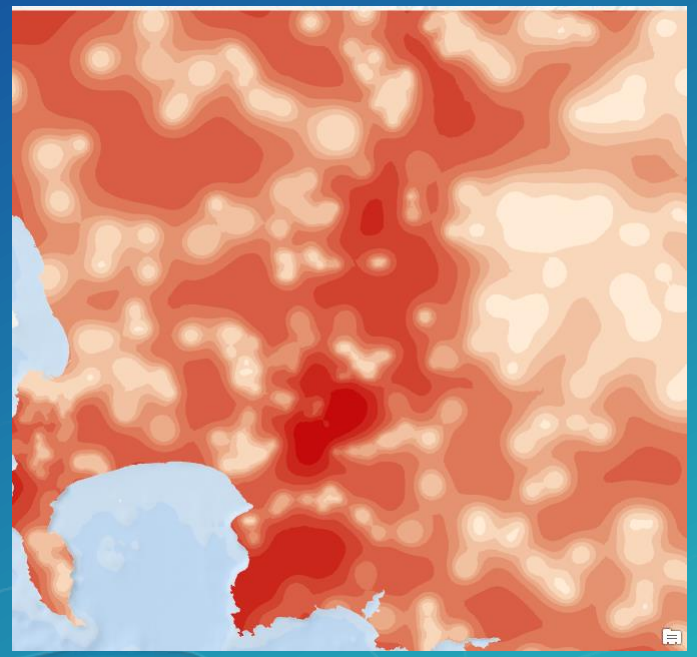


Output surfaces

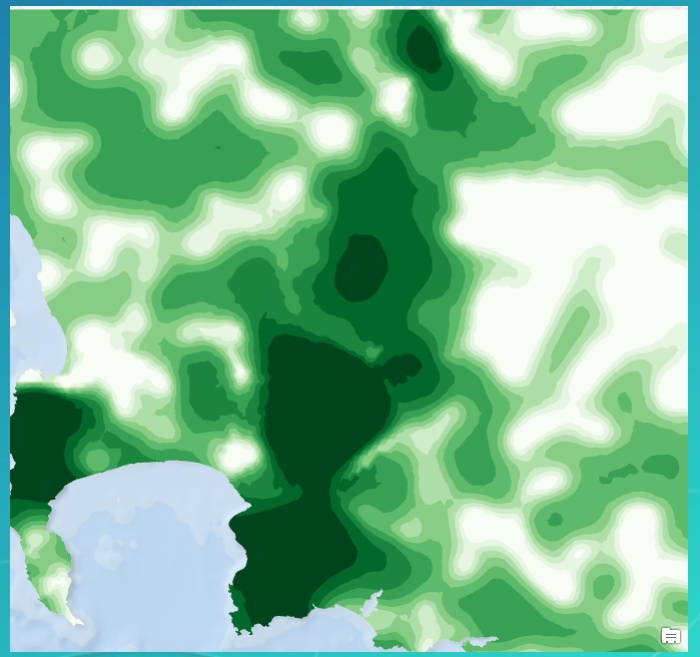
Prediction



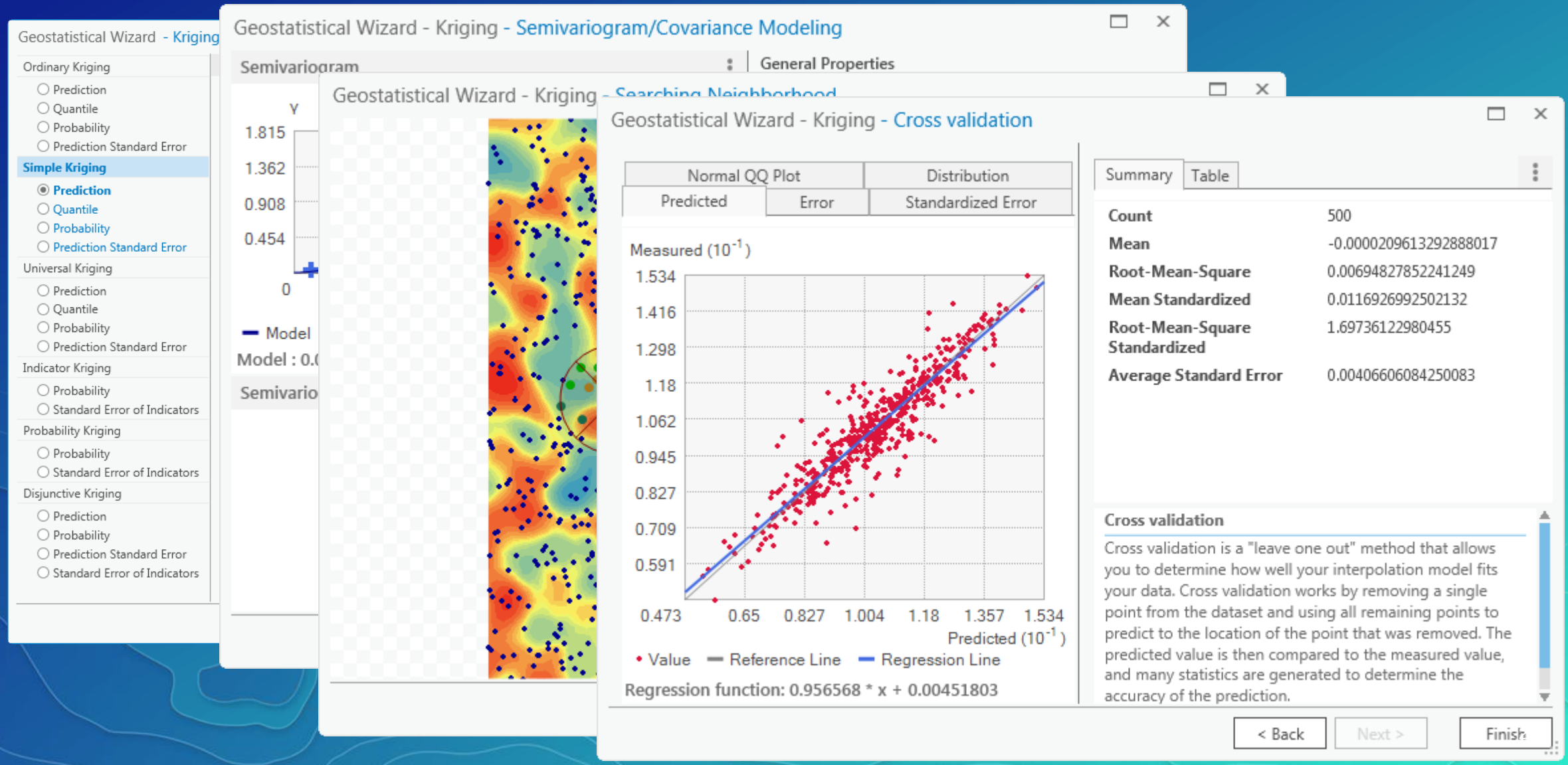
Standard error of prediction



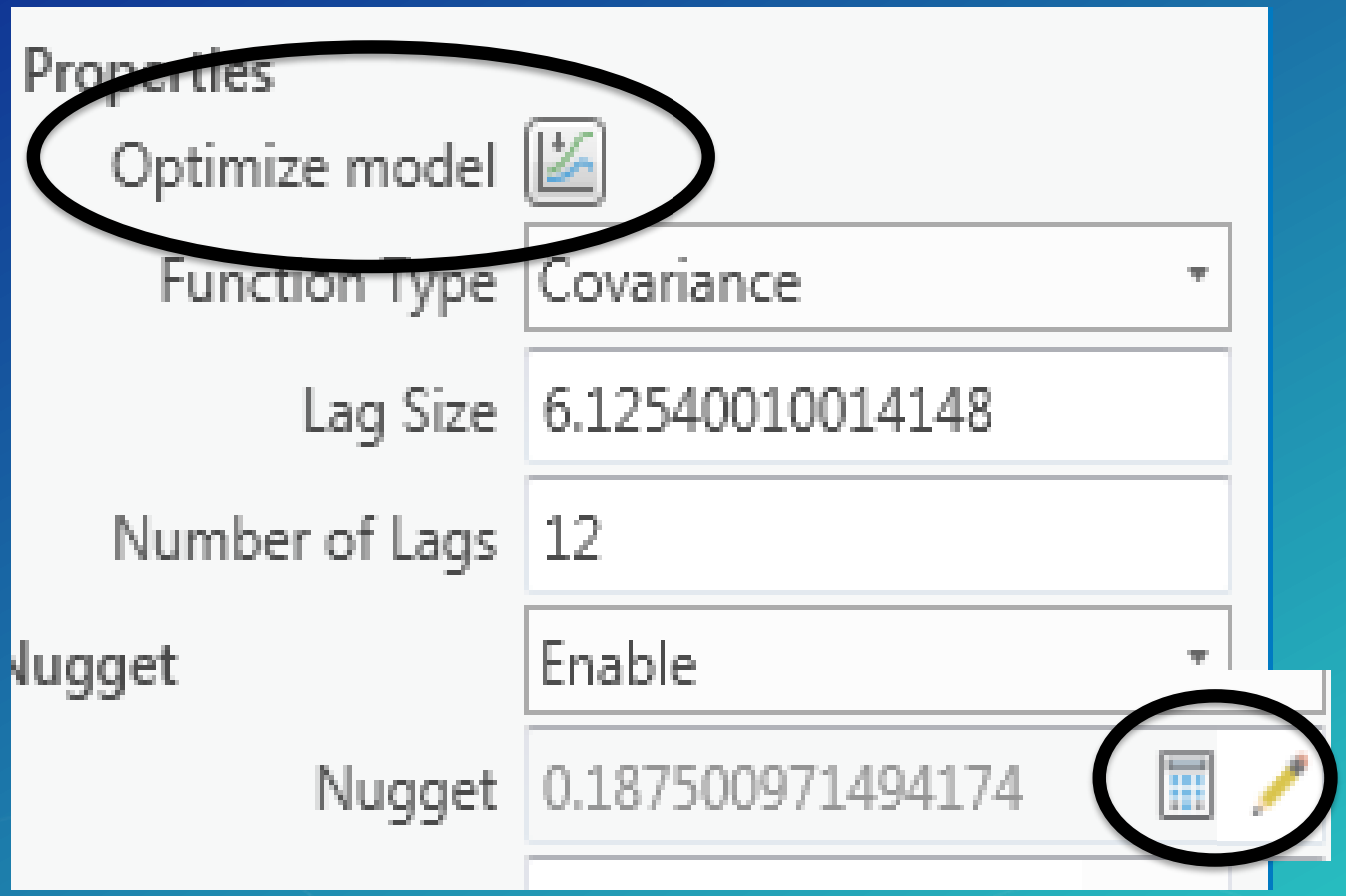
Probability that rainfall exceeds 900mm



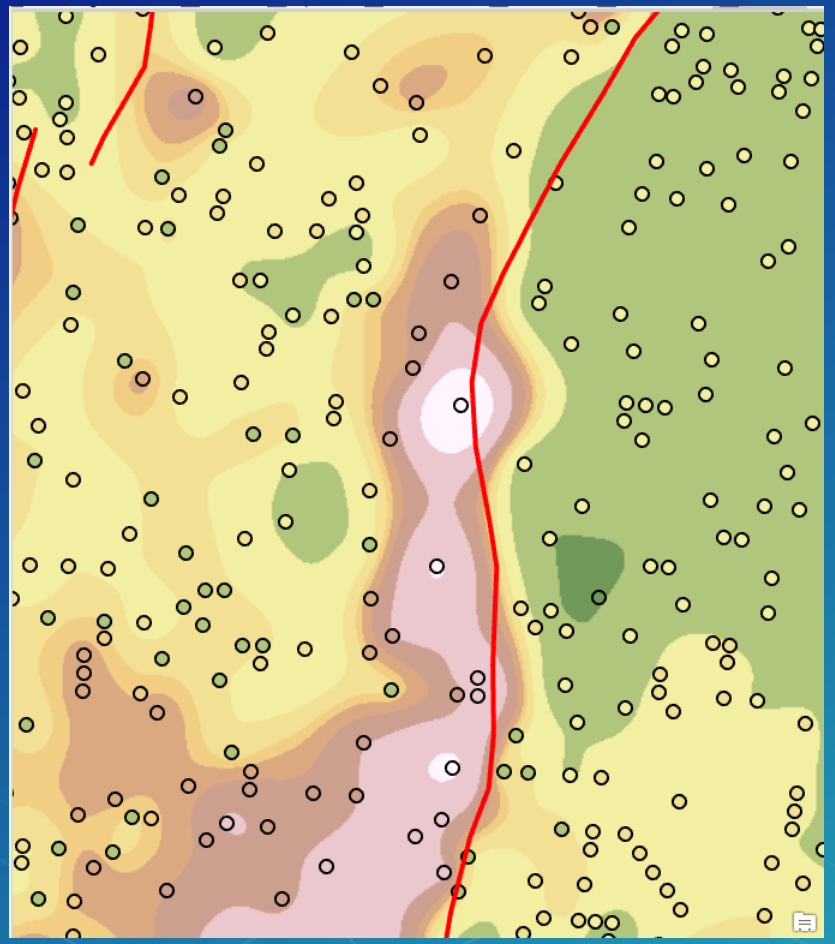
Interactive Wizard

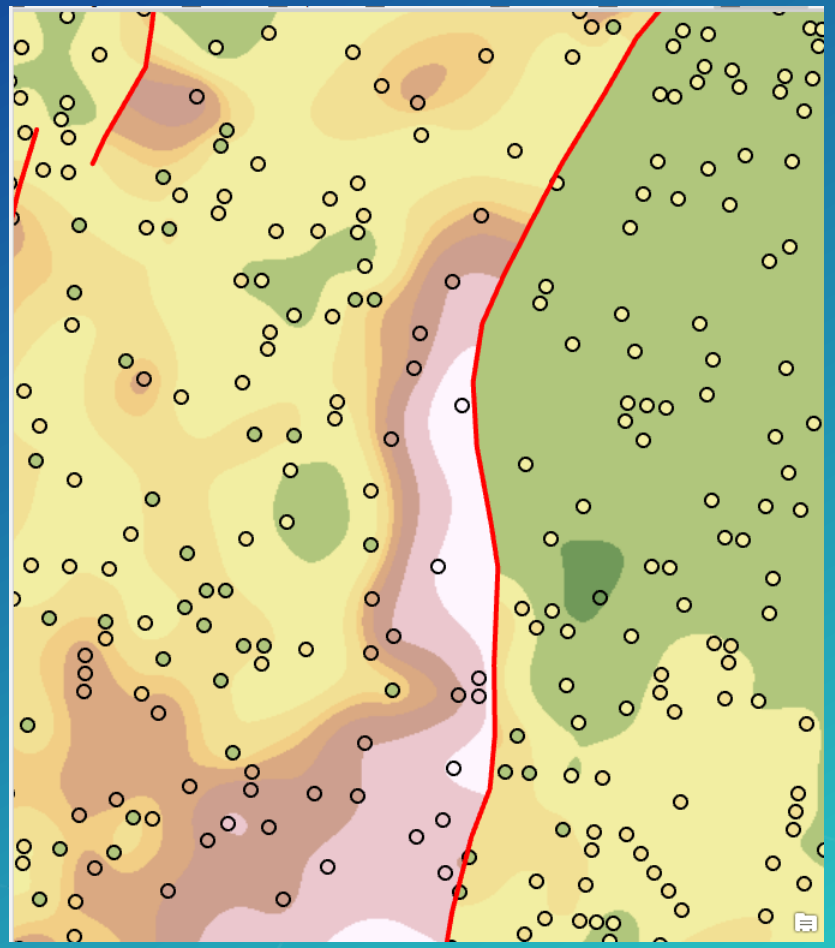


Interactive Wizard



Barriers

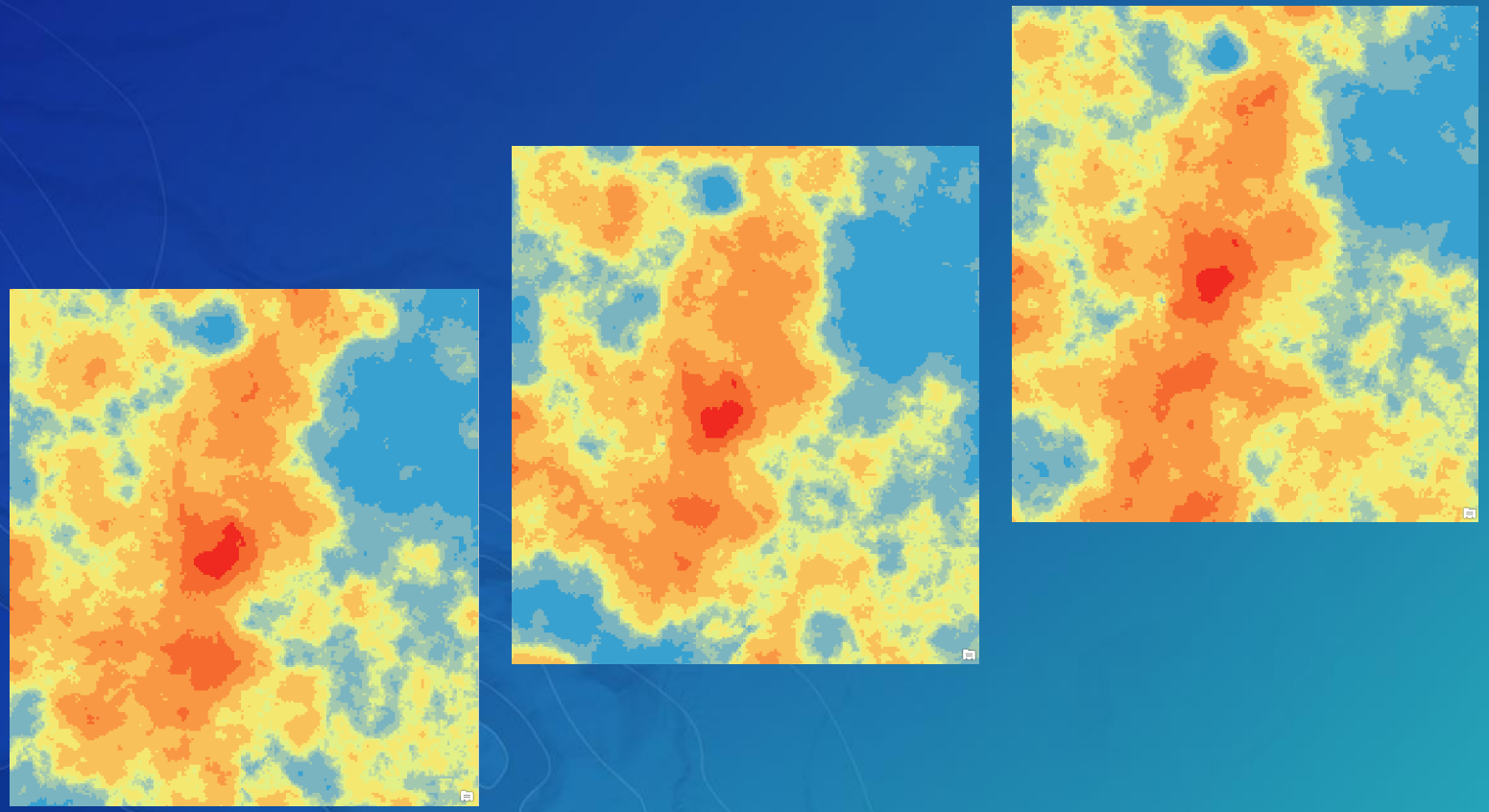


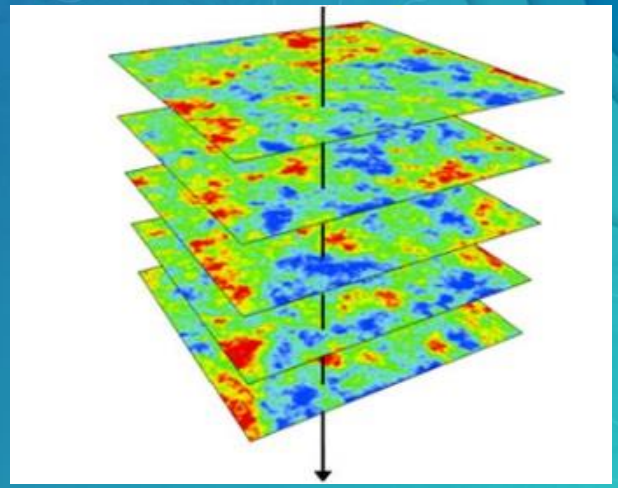


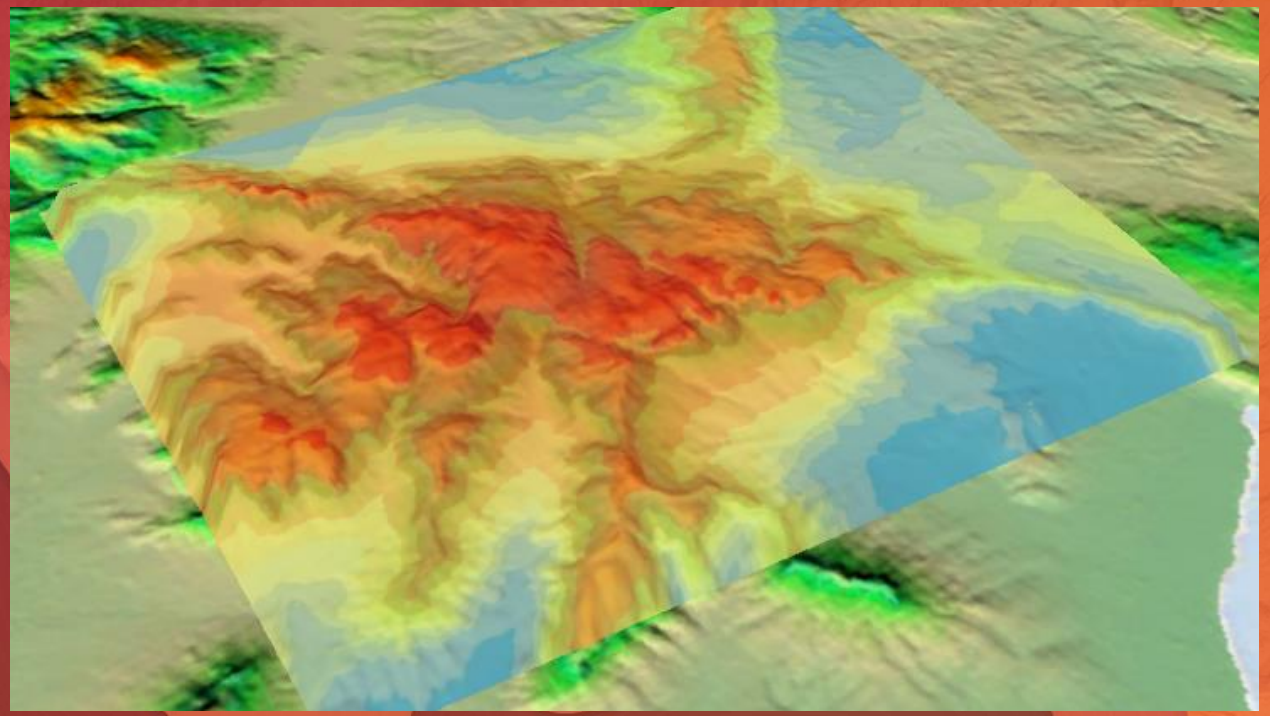
Gaussian Geostatistical Simulations

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

- Any realization might be the “real” thing!

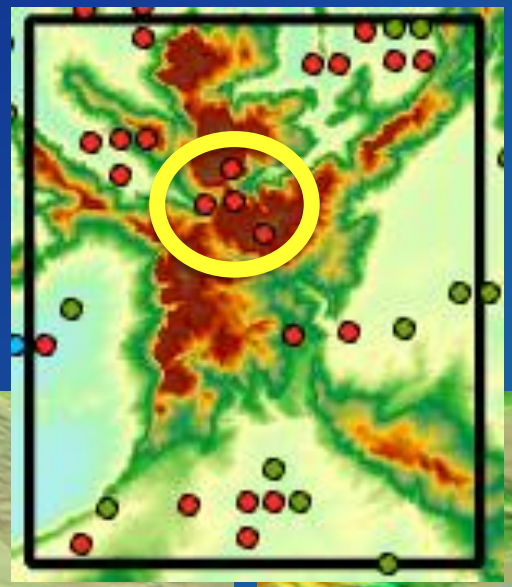


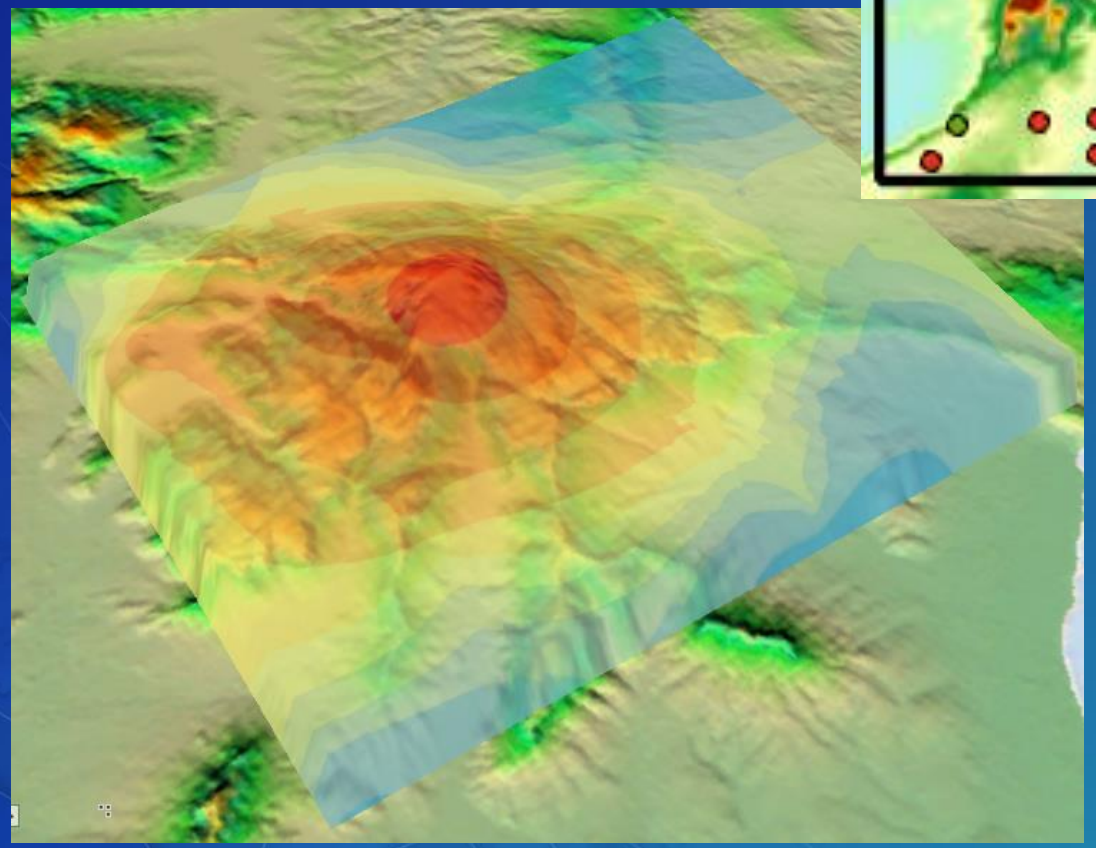


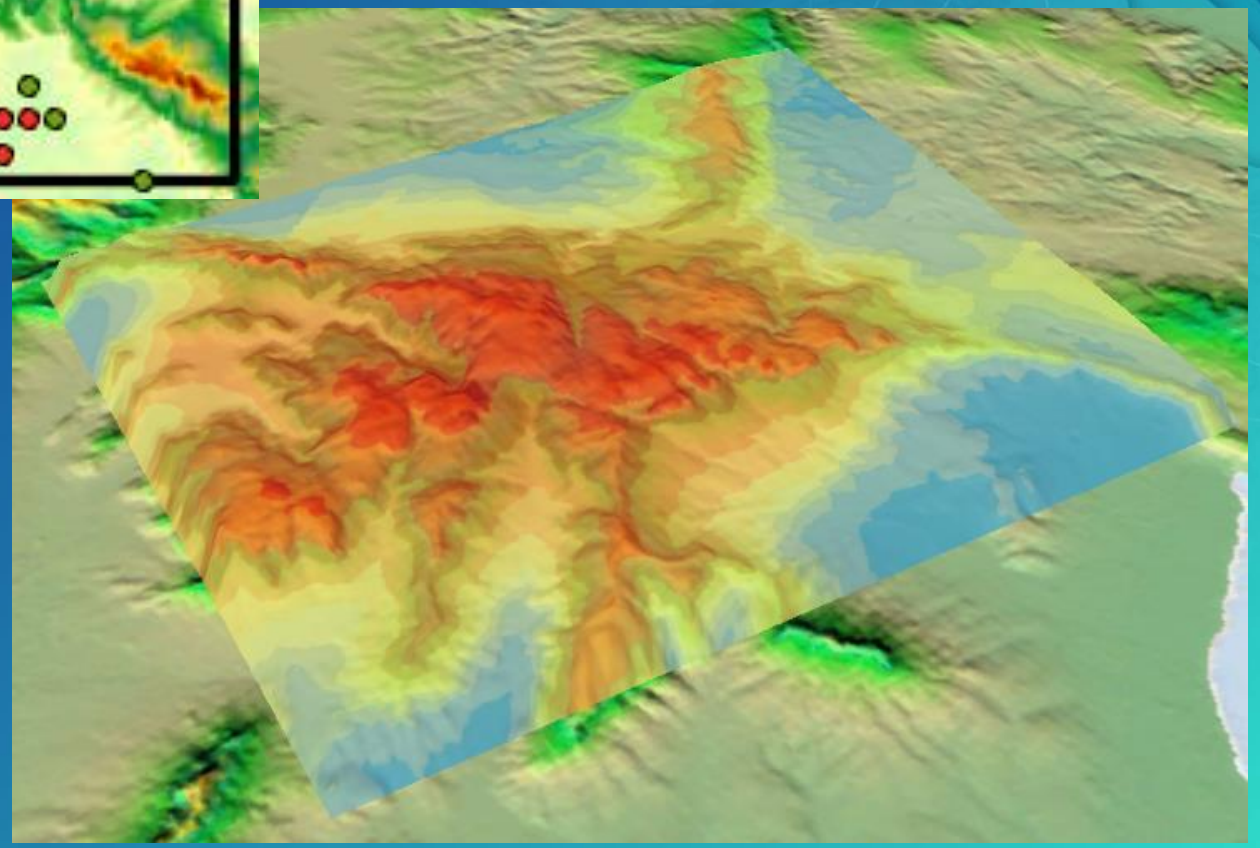


EBK Regression Prediction

Eric Krause







EBK Regression Prediction

Geostatistical layers

Eric Krause

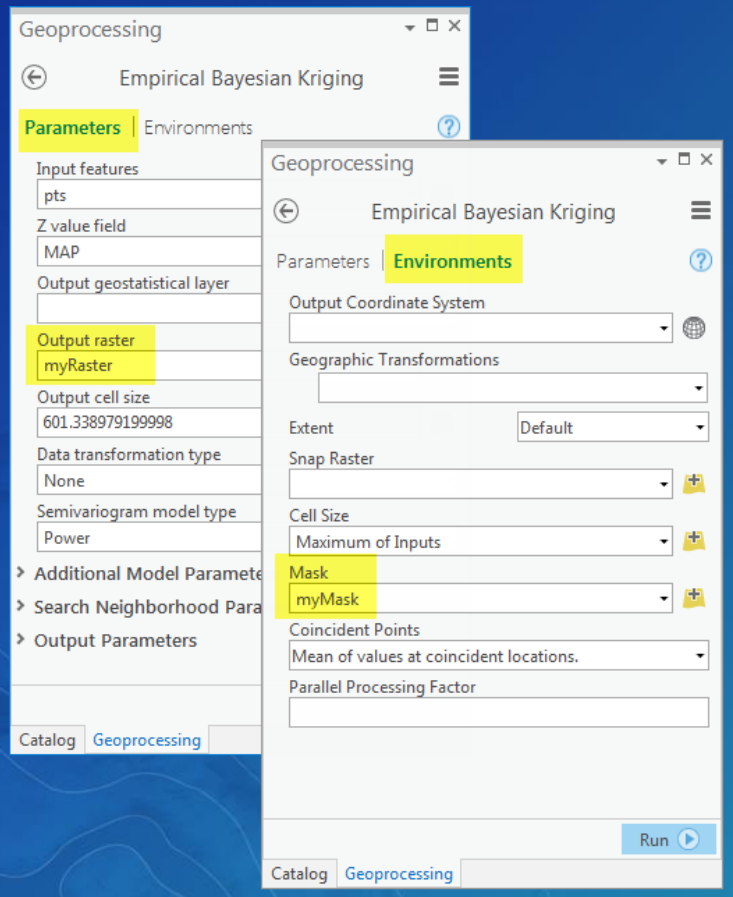
Where is GA used?

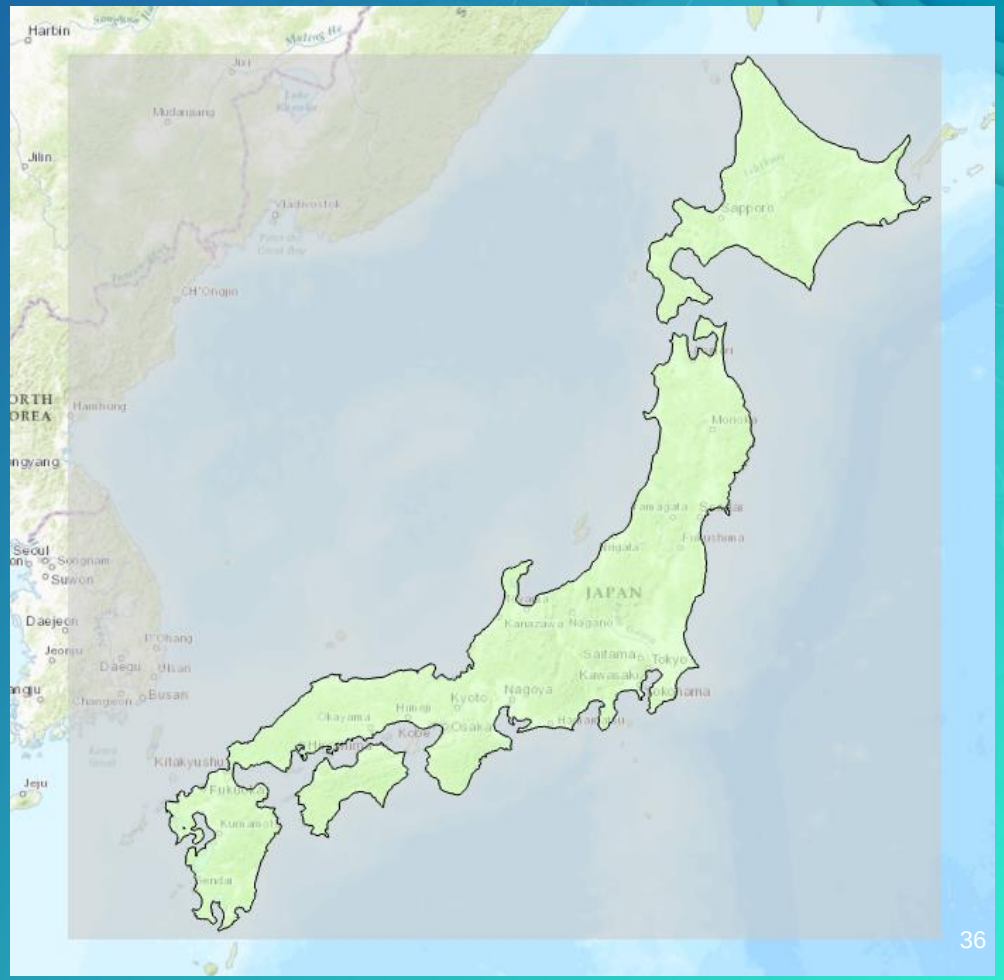
- Anyone who needs to statistically explore data and create surfaces for a number of variables will benefit from this statistical software package.

- 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

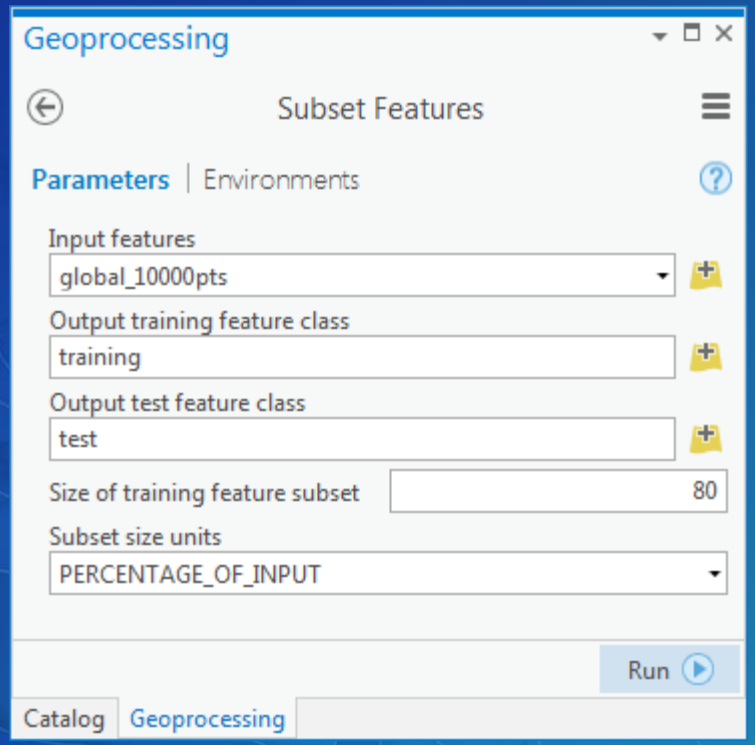
- Use Mask when creating a raster
  - 8700 pixels inside (55,000 outside) Japan (6 ½ times)

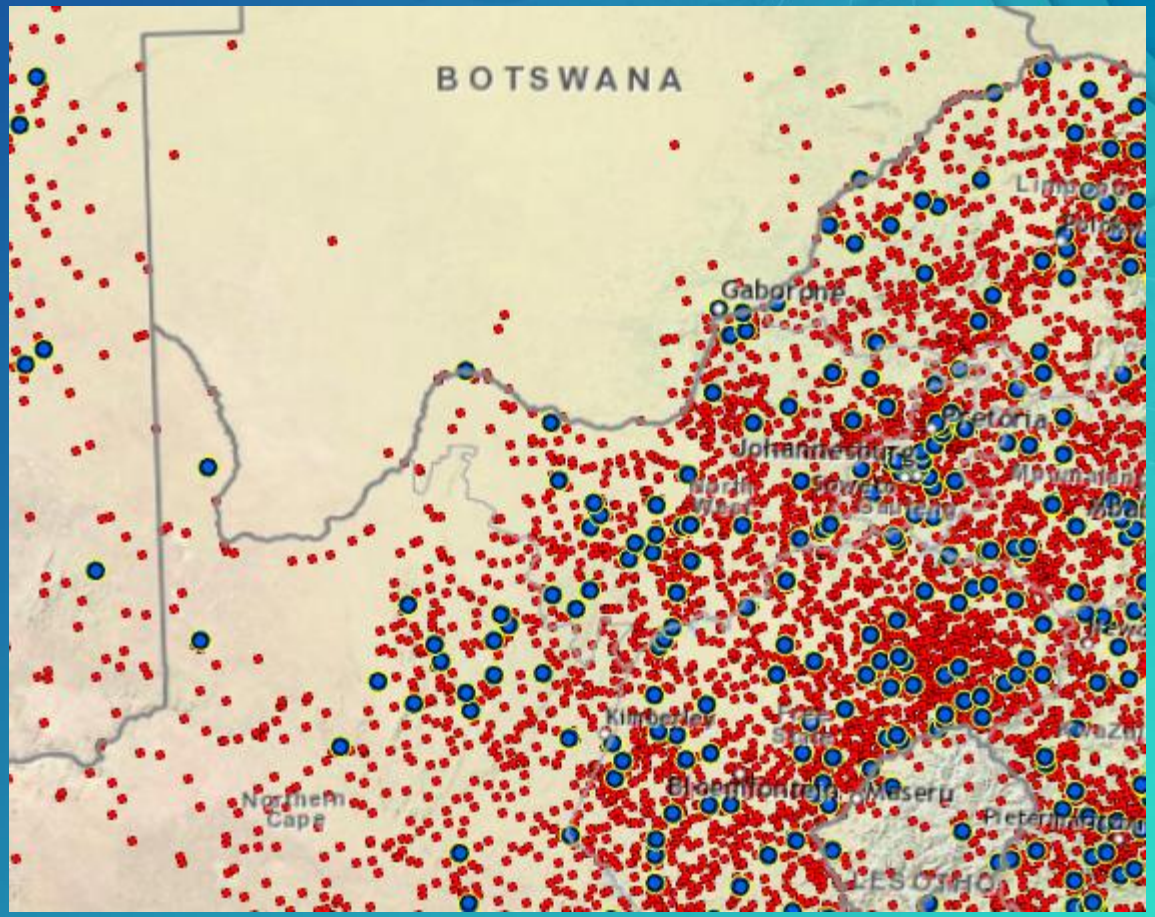




Tips & Tricks

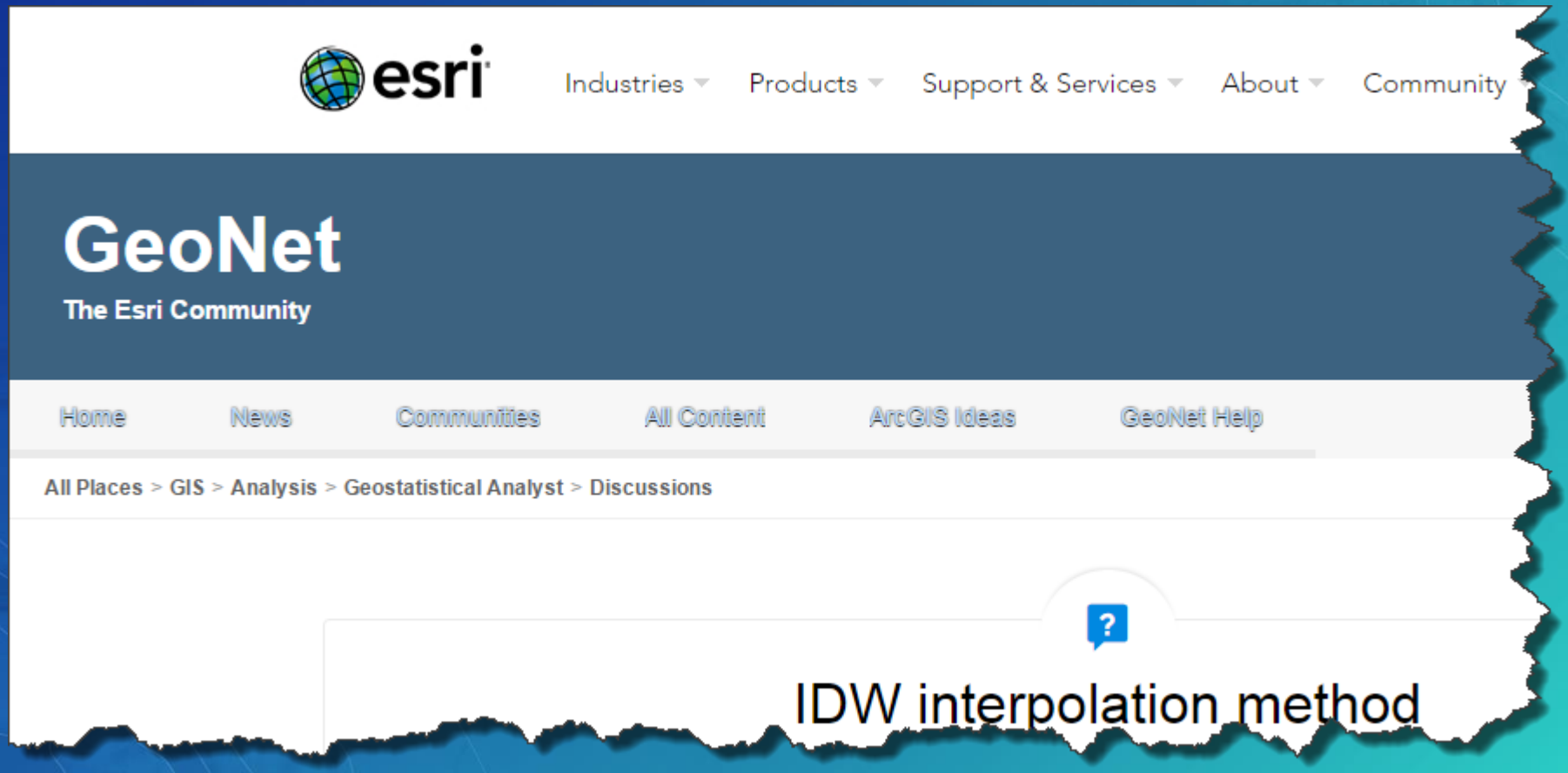
- Subset of the data
  - SubsetFeatures GP tool
  - Selection





Conclusions

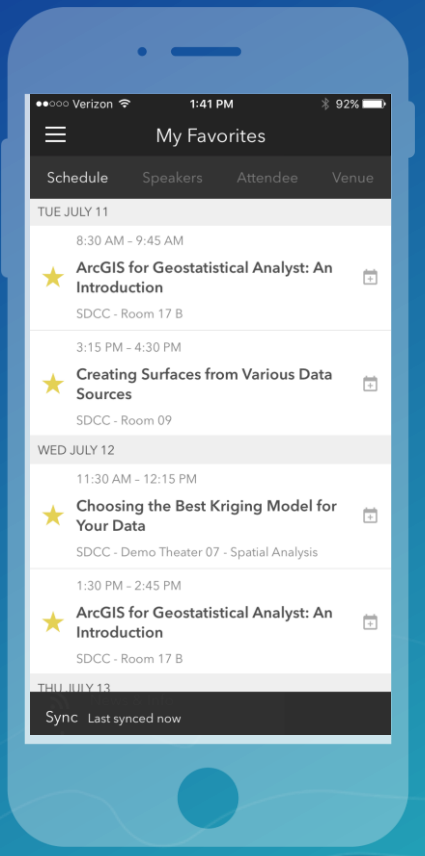
https://geonet.esri.com/



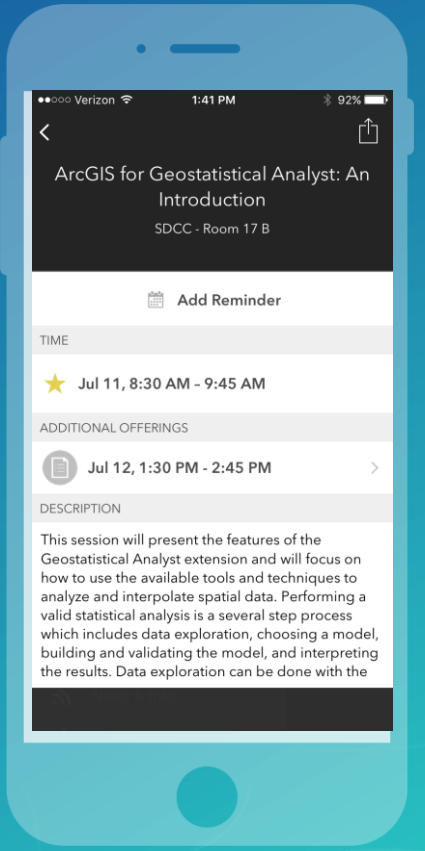
Please Take Our Survey on the **Esri Events App!**

Download the Esri Events app and find your event

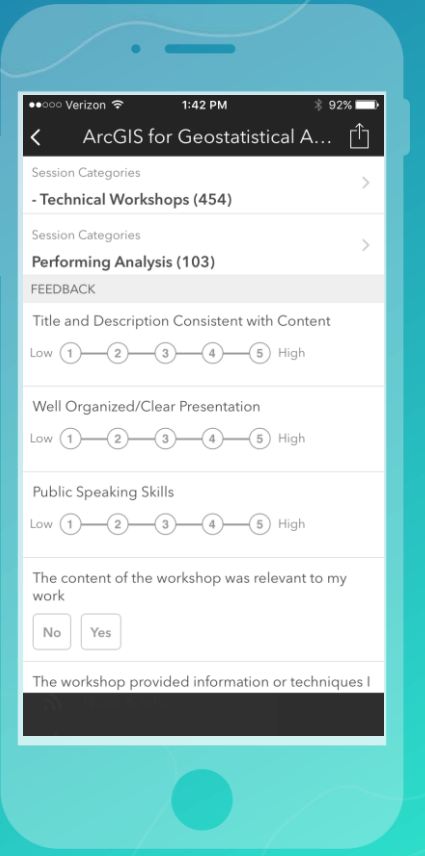
Select the session you attended



Scroll down to find the survey



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())
```