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 $h$) = $0.5 \times \text{average} \left[ (\text{value}_i - \text{value}_j)^2 \right]$

- **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
  - 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)
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 - Empirical Bayesian Kriging

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
Unlike smoothing the output, this method modifies the weights.
Search neighborhood - Standard

2 per sector

8 closest
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

![Geostatistical Wizard - Kriging - Cross validation](image)

<table>
<thead>
<tr>
<th>Source</th>
<th>Measured</th>
<th>Predicted</th>
<th>Error</th>
<th>Standard Error</th>
<th>Standardized Error</th>
<th>Normal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>428.299987</td>
<td>461.0831795</td>
<td>32.78319178</td>
<td>158.822240</td>
<td>0.20652448013545</td>
<td>0.402527890259</td>
</tr>
<tr>
<td>2</td>
<td>89.93792032</td>
<td>159.606806</td>
<td>69.66888552</td>
<td>0.56392492919584</td>
<td>-1.468527578576</td>
<td>-0.0796508797952</td>
</tr>
<tr>
<td>3</td>
<td>3.720764510</td>
<td>157.3111541</td>
<td>0.02365219199303</td>
<td>-0.719908259752</td>
<td>1.4493019584621</td>
<td>1.4493019584621</td>
</tr>
<tr>
<td>4</td>
<td>104.2607293</td>
<td>157.839016</td>
<td>0.0610516665327</td>
<td>0.66055106665327</td>
<td>0.5643075659010</td>
<td>0.5643075659010</td>
</tr>
<tr>
<td>5</td>
<td>12.311307311</td>
<td>156.356398</td>
<td>0.2706080958049</td>
<td>0.5643075659010</td>
<td>0.5643075659010</td>
<td>0.5643075659010</td>
</tr>
<tr>
<td>6</td>
<td>84.055296227</td>
<td>159.962529</td>
<td>-0.5256686828154</td>
<td>-1.423936348983</td>
<td>-1.423936348983</td>
<td>-1.423936348983</td>
</tr>
<tr>
<td>7</td>
<td>91.61388226</td>
<td>156.753627</td>
<td>-0.5844418660042</td>
<td>-1.499322001875</td>
<td>-1.499322001875</td>
<td>-1.499322001875</td>
</tr>
<tr>
<td>8</td>
<td>77.7558023717596</td>
<td>10.6016296</td>
<td>66.6640736706294</td>
<td>1.461466417715</td>
<td>1.461466417715</td>
<td>1.461466417715</td>
</tr>
<tr>
<td>9</td>
<td>8.0137622990752729</td>
<td>18.36437298</td>
<td>-0.1198626313017</td>
<td>0.345329320507</td>
<td>0.345329320507</td>
<td>0.345329320507</td>
</tr>
<tr>
<td>10</td>
<td>38.2520385</td>
<td>168.021862</td>
<td>0.8228217247287</td>
<td>1.746876871671</td>
<td>1.746876871671</td>
<td>1.746876871671</td>
</tr>
<tr>
<td>11</td>
<td>71.13504292</td>
<td>162.7596</td>
<td>0.43706107387348</td>
<td>0.966459776804</td>
<td>0.966459776804</td>
<td>0.966459776804</td>
</tr>
<tr>
<td>12</td>
<td>82.96394155251061</td>
<td>20.38557130</td>
<td>-0.128502647737</td>
<td>-0.576361702568</td>
<td>-0.576361702568</td>
<td>-0.576361702568</td>
</tr>
</tbody>
</table>

Count: 8319
Average CRPS: 36.8364340137905
Inside 90 Percent Interval: 91.1047000041447
Inside 95 Percent Interval: 94.951315263974
Mean: 1.26256965544221
Root-Mean-Square: 77.7558023717596
Mean Standardized: 0.0137622990752729
Root-Mean-Square Standardized: 0.961343742125846
Average Standard Error: 8.26394155251061

Regression function: 1.04431 * x - 34.8993
Output surfaces

Prediction

Standard error of prediction

Probability that rainfall exceeds 900mm
Interactive Wizard

Geostatistical Wizard - Kriging - Semivariogram/Covariance Modeling

Geostatistical Wizard - Kriging - Cross validation

Summary

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>500</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.003029531228388017</td>
</tr>
<tr>
<td>Root-Mean-Square</td>
<td>0.0094987552214249</td>
</tr>
<tr>
<td>Mean Standardized</td>
<td>0.011692592502132</td>
</tr>
<tr>
<td>Root-Mean-Square Standardized</td>
<td>1.69736122980455</td>
</tr>
<tr>
<td>Average Standard Error</td>
<td>0.00406606082450083</td>
</tr>
</tbody>
</table>

Cross validation is a "leave one out" method that allows you to determine how well your interpolation model fits your data. Cross validation works by removing a single point from the dataset and using all remaining points to predict the location of the point that was removed. The predicted value is then compared to the measured value, and many statistics are generated to determine the accuracy of the prediction.
Interactive Wizard

Properties

- Optimize model

Function Type: Covariance

Lag Size: 6.12540010014148

Number of Lags: 12

Nugget: Enable

Nugget: 0.187500971494174

Buttons:
- < Back
- Next >
- Finish
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/

IDW interpolation method
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.”
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(inPC, 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())
```

http://esriurl.com/GeostatGetStarted