Empirical Bayesian Kriging and EBK Regression Prediction in ArcGIS

Eric Krause

Sessions of note…

Tuesday

- Interpolating Surfaces in ArcGIS (1:00-2:00 SDCC Rm33C)
- Kriging: An Introduction to Concepts and Applications (2:30-3:30 SDCC Rm33C)
- Geostatistical Analyst: An Introduction (4:00-5:00 SDCC Rm30C)

Wednesday

- Surface Interpolation in ArcGIS (11:15-12:00 SDCC Demo Theater 10)
- Empirical Bayesian Kriging and EBK Regression Prediction in ArcGIS (2:30-3:15 SDCC Demo Theater 10)

Thursday

- Geostatistics in Practice: Learning Interpolation Through Examples (8:30-9:30 SDCC Rm30A)
- Polygon-to-Polygon Predictions Using Areal Interpolation (11:15-12:00 SDCC Demo Theater 10)
- Geostatistical Analyst: An Introduction (1:00-2:00 SDCC Rm30A)
- Using Living Atlas Data and ArcGIS Pro for 3D Interpolation (2:30-3:30 SDCC Rm 30C)
- Interpolating Surfaces in ArcGIS (4:00-5:00 SDCC Rm15A)
- Kriging: An Introduction to Concepts and Applications (4:00-5:00 SDCC Rm15B)

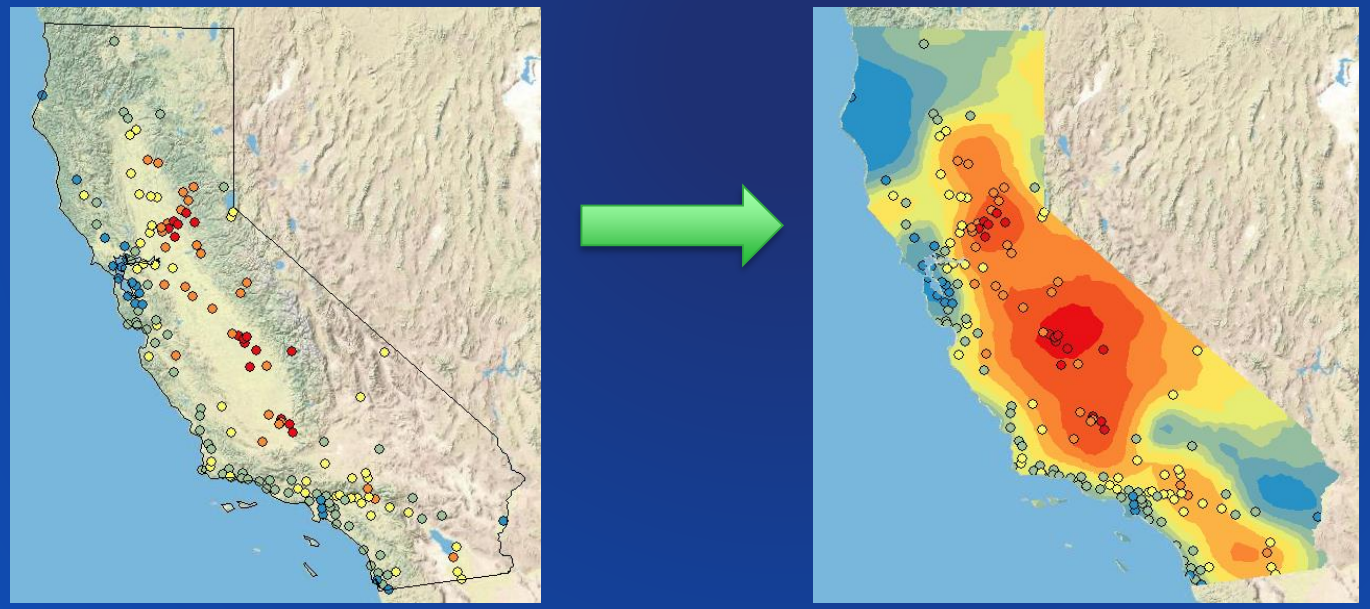
Geostatistical Analyst Resources

http://esriurl.com/GeostatGetStarted

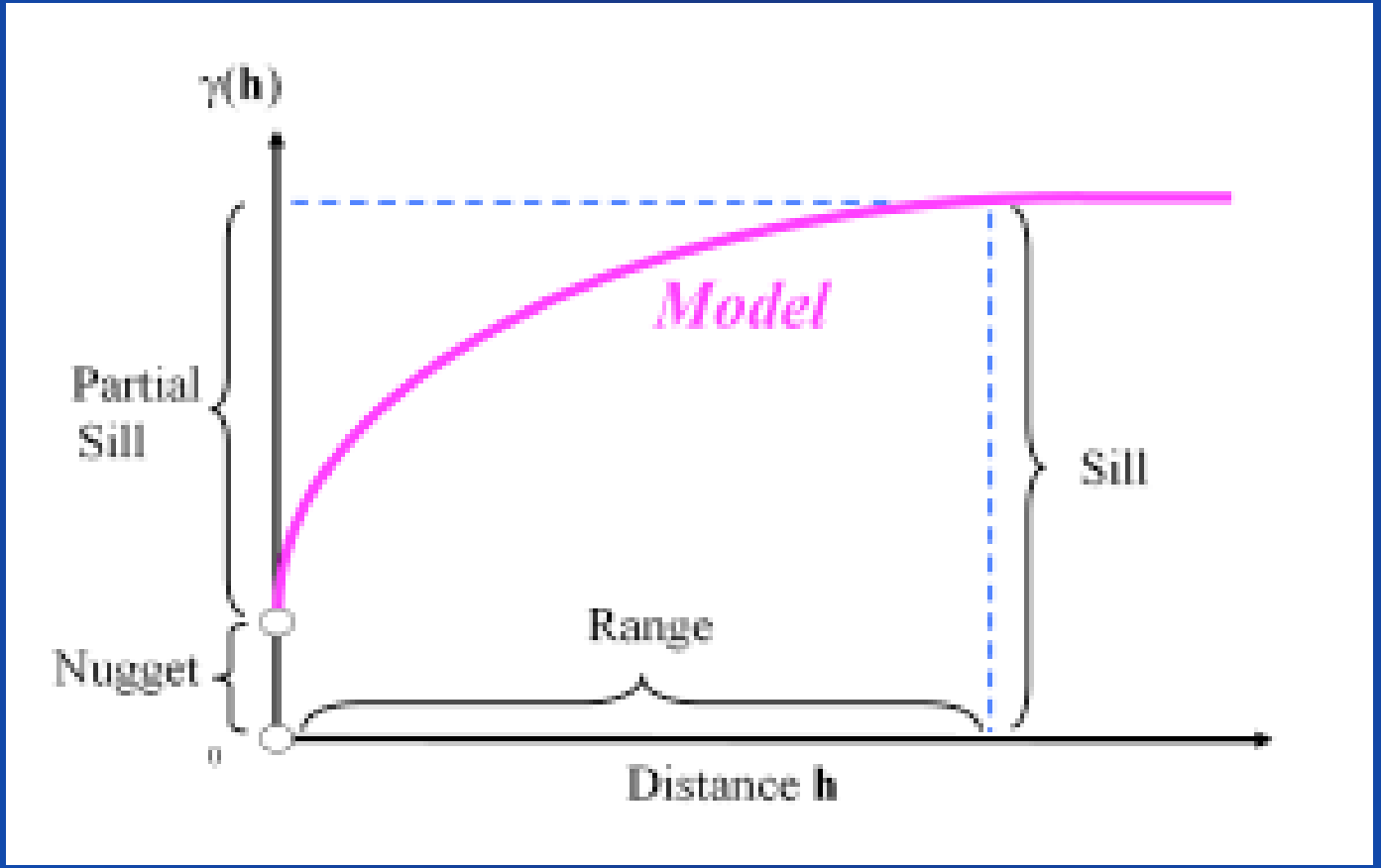
- GeoNet – community.esri.com
  - Blogs
  - Free textbook and datasets
    - Spatial Statistical Analysis For GIS Users
  - Lots of discussions and Q&A
- Learn GIS – learn.arcgis.com
  - Model Water Quality Using Interpolation
  - Analyze Urban Heat Using Kriging
  - Interpolate 3D Oxygen Measurements in Monterey Bay

What is interpolation?

- Predict values at unknown locations using values at measured locations
- Many interpolation methods: kriging, IDW, LPI, etc



Semivariogram Modeling



Empirical Bayesian Kriging

- Advantages
  - Requires minimal interactive modeling, spatial relationships are modeled automatically
  - Usually more accurate, especially for small or nonstationary datasets
  - Uses local models to capture small scale effects
    - Doesn’t assume one model fits the entire data
  - Standard errors of prediction are more accurate than other kriging methods
- Disadvantages
  - Processing is slower than other kriging methods
  - Limited customization

How does EBK work?

1. Divide the data into subsets of a given size
   - Controlled by “Subset Size” parameter
   - Subsets can overlap, controlled by “Overlap Factor”
2. For each subset, estimate the semivariogram
3. Simulate data at input point locations and estimate new semivariogram from the simulated data
4. Repeat step 3 many times. This results in a distribution of semivariograms
   - Controlled by “Number of Simulations”
5. Mix the local surfaces together to get the final surface.

EBK Regression Prediction

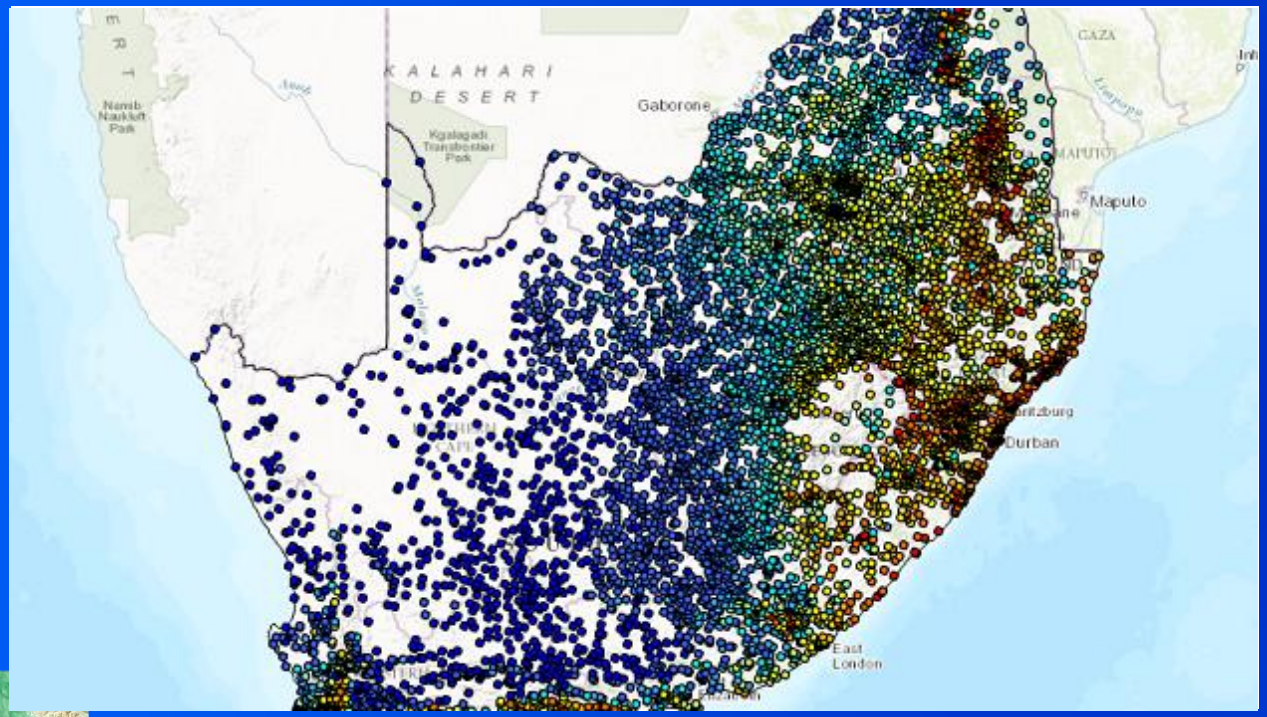
- Allows you to use explanatory variable rasters to improve predictions
- Automatically extracts useful information from explanatory variables
- Uses Principle Components to handle multicollinearity

Transformations

- **Two available transformations**
  - **Empirical – Fits a smooth distribution to the data, then transforms to normal distribution. Useful for data that is not bell-shaped**
  - **Log Empirical – Takes logarithm of data before performing Empirical transformation. Useful for data that cannot be negative (eg, rainfall)**

Data in Geographic Coordinate Systems

- Euclidean distance for geographic coordinates is very inaccurate, particularly far from the equator
- EBK uses chordal distances
  - Chordal distance is the 3D straight-line distance between points on a spheroid
  - Accurate approximation to geodesic distance up to 30 degrees



Demo

Eric Krause


esri
THE
SCIENCE
OF
WHERE