Concepts and Applications of Kriging

Eric Krause

Konstantin Krivoruchko

Sessions of note…

Tuesday

- *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

- *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)

Outline

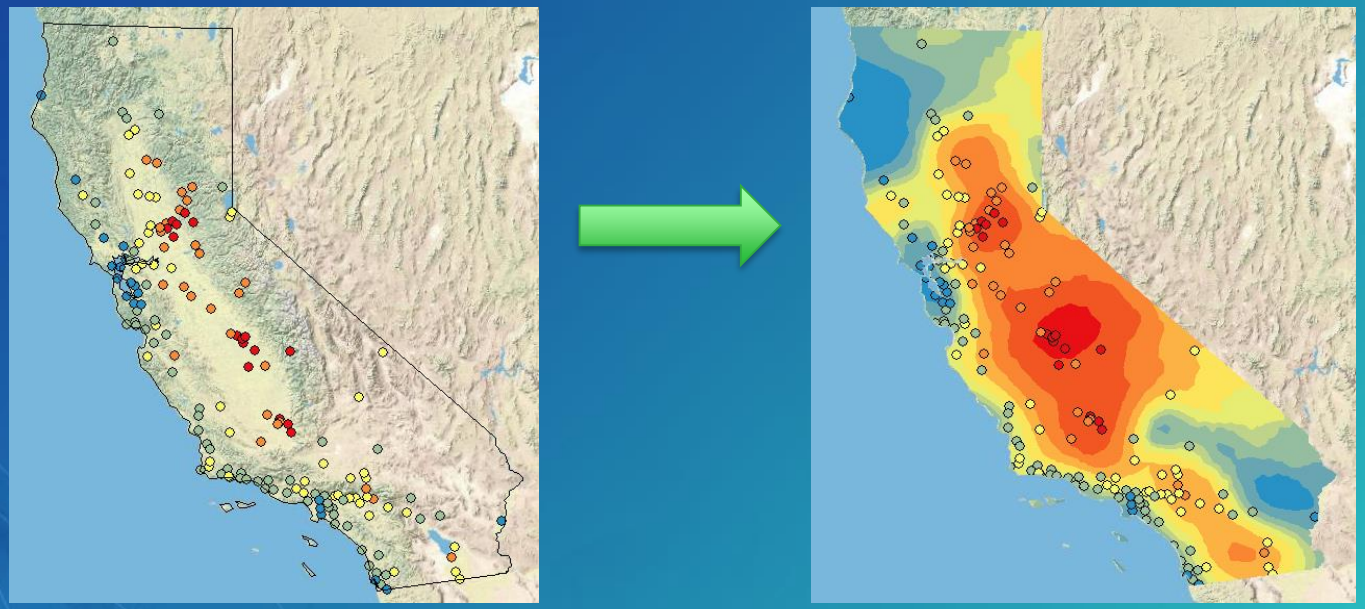
- Introduction to interpolation
- What is new in Geostatistical Analyst?
- Introduction to kriging
- Validating interpolation results
- Empirical Bayesian Kriging and EBK Regression Prediction
- Areal Interpolation
- Questions

What is new in Geostatistical Analyst?

- Pro, Pro, PRO! It's the future, get on board.
- In ArcGIS Pro 1.4:
  - Geostatistical Wizard
- In ArcGIS Pro 2.0:
  - Areal Interpolation
  - Interactive cross validation and model comparison
  - Directional semivariograms
- In the future (maybe)…
  - Explanatory regression
  - More charting tools
  - 3D and Space-Time interpolation

What is interpolation?

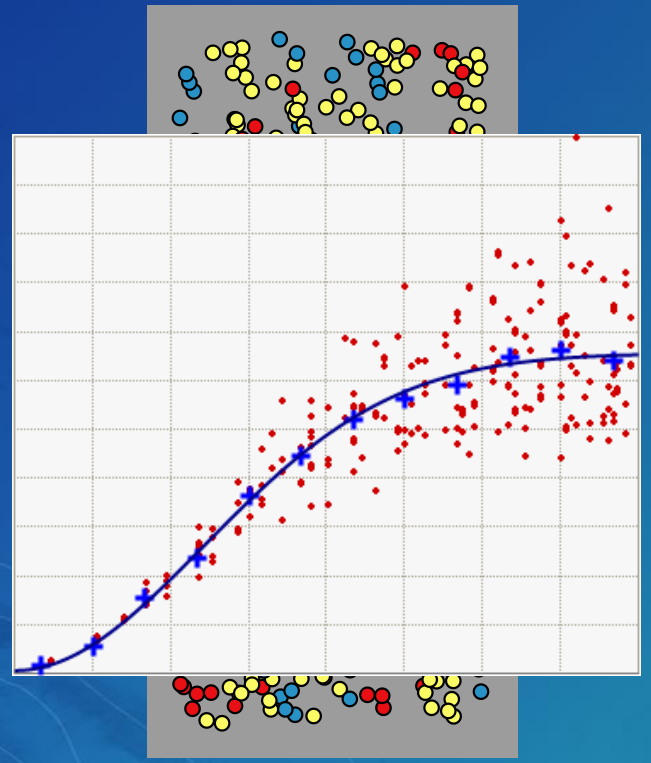
- Predict values at unknown locations using values at measured locations
- Many interpolation methods: kriging, IDW, LPI, etc

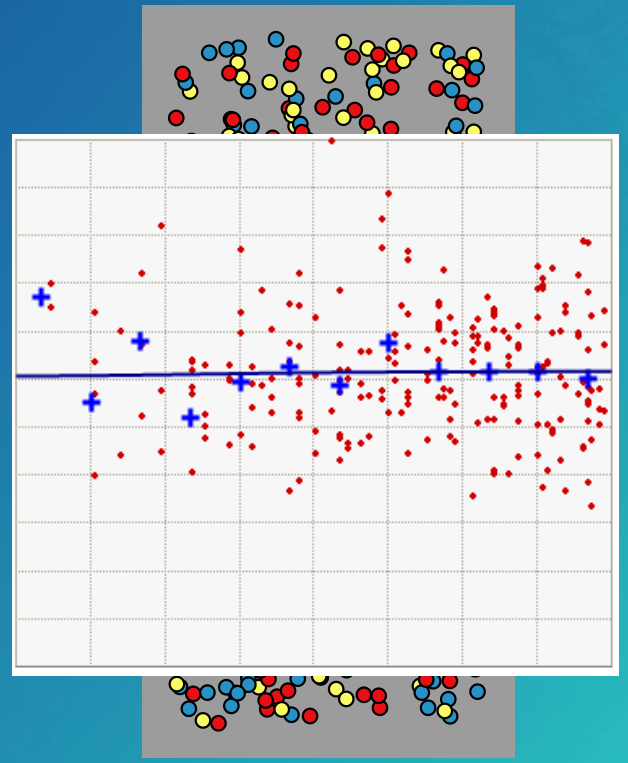


What is autocorrelation?

Tobler’s first law of geography:

"Everything is related to everything else,
but near things are more related than distant things."





What is kriging?

- Kriging is the optimal interpolation method if the data meets certain conditions.
- What are these conditions?
  - Normally distributed
  - Stationary
  - No trends
- How do I check these conditions?
  - Exploratory Analysis and charting

What is an “optimal” interpolator?

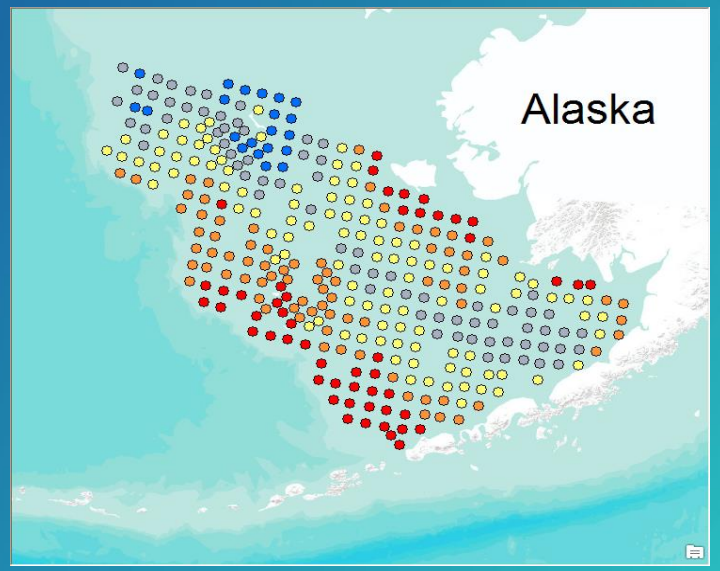
- Estimates the true value, on average
- Lowest expected prediction error
- Able to use extra information, such as covariates
- Filters measurement error
- Can be generalized to polygons (Areal interpolation, Geostatistical simulations)
- Can estimate quantiles to test best/worst case scenarios

Geostatistical workflow

1. Explore the data
2. Choose an interpolation method
3. Fit the interpolation model
4. Validate the results
5. Repeat steps 2-4 as necessary
6. Map the data for decision-making

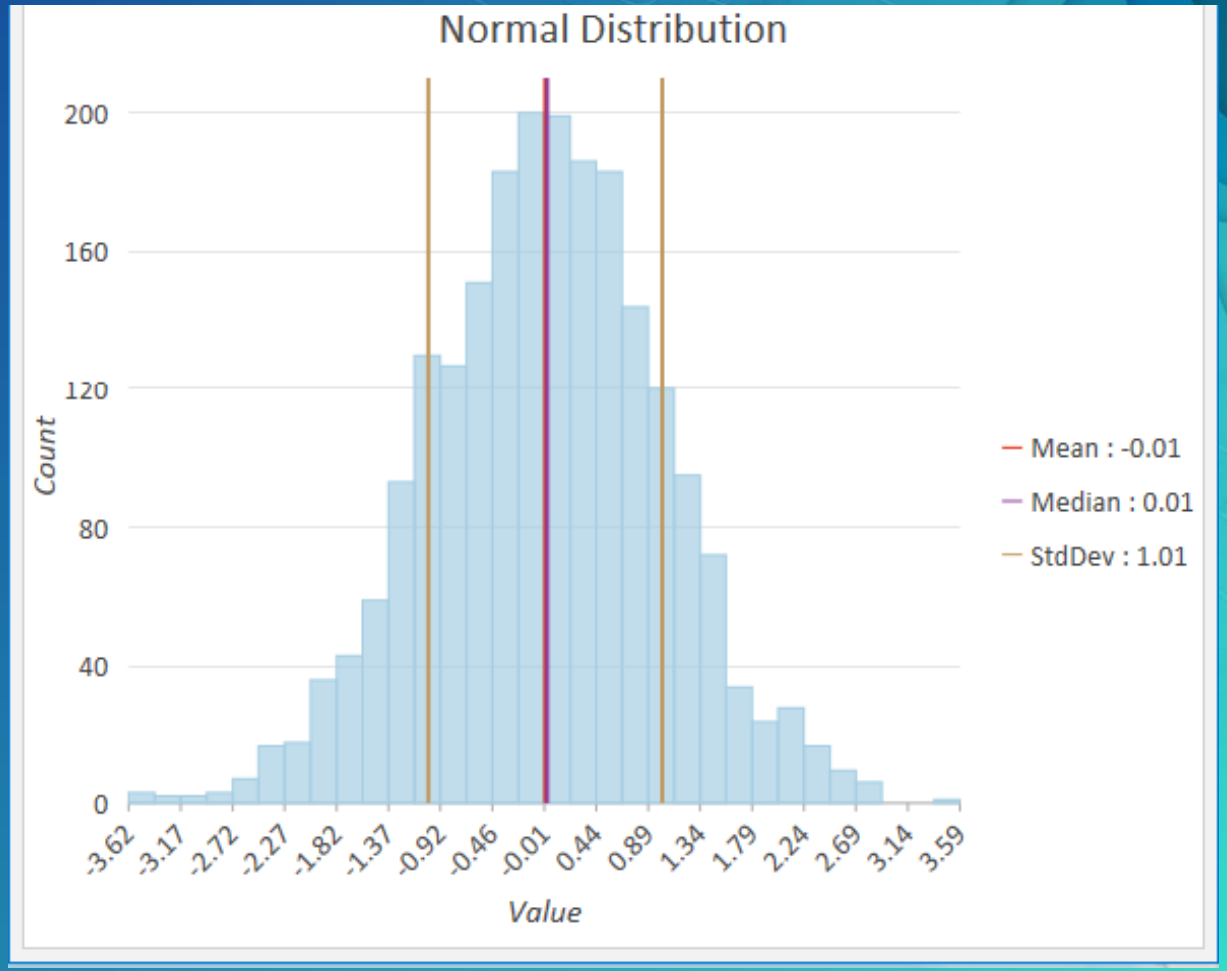
Exploring your data

1. Where is the data located?
2. What are the values of the data points?
3. How does the location of a point relate to its value?

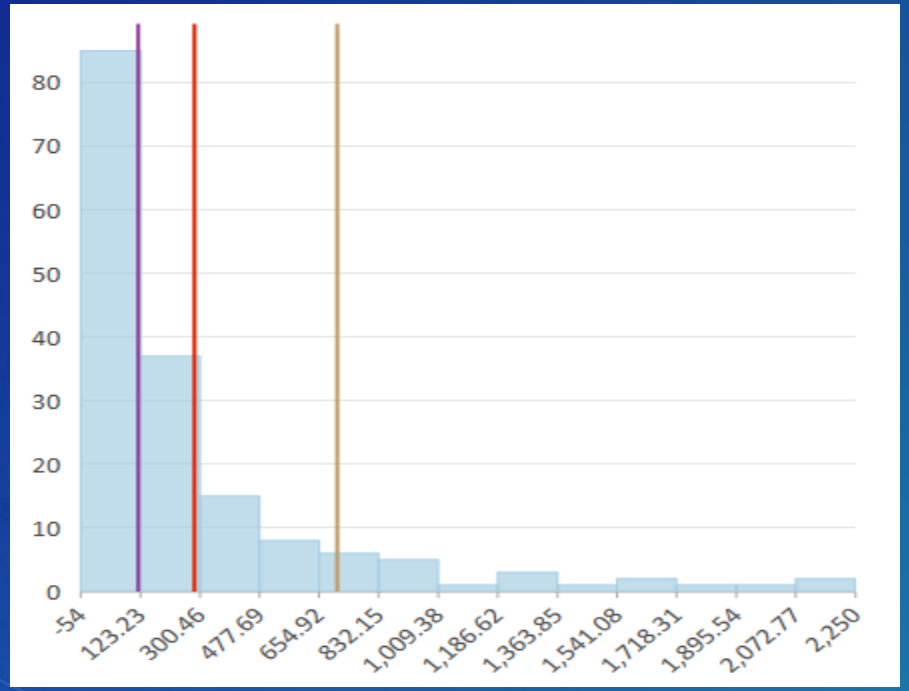


Does my data follow a normal distribution?

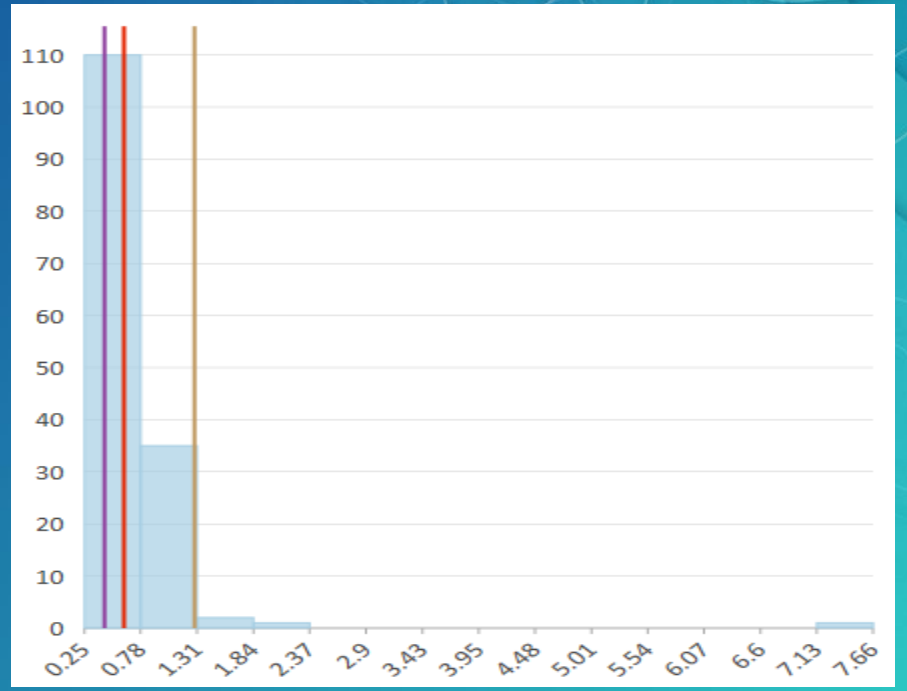
- **Histogram chart**
  - Symmetric and bell-shaped
  - Look for outliers
  - Mean ≈ Median
- **What can I do if my data is not normally distributed?**
  - Apply a transformation
    - Log, Box Cox, Arcsin, Normal Score Transformation



Skewed distributions and outliers



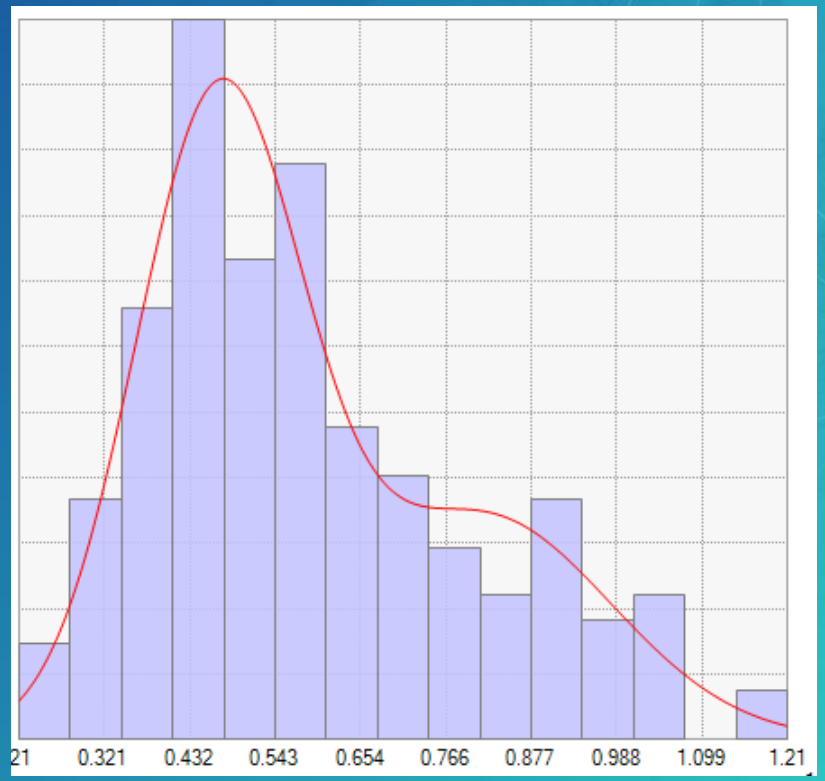
Skewed Distribution



Outlier

Normal Score Transformation

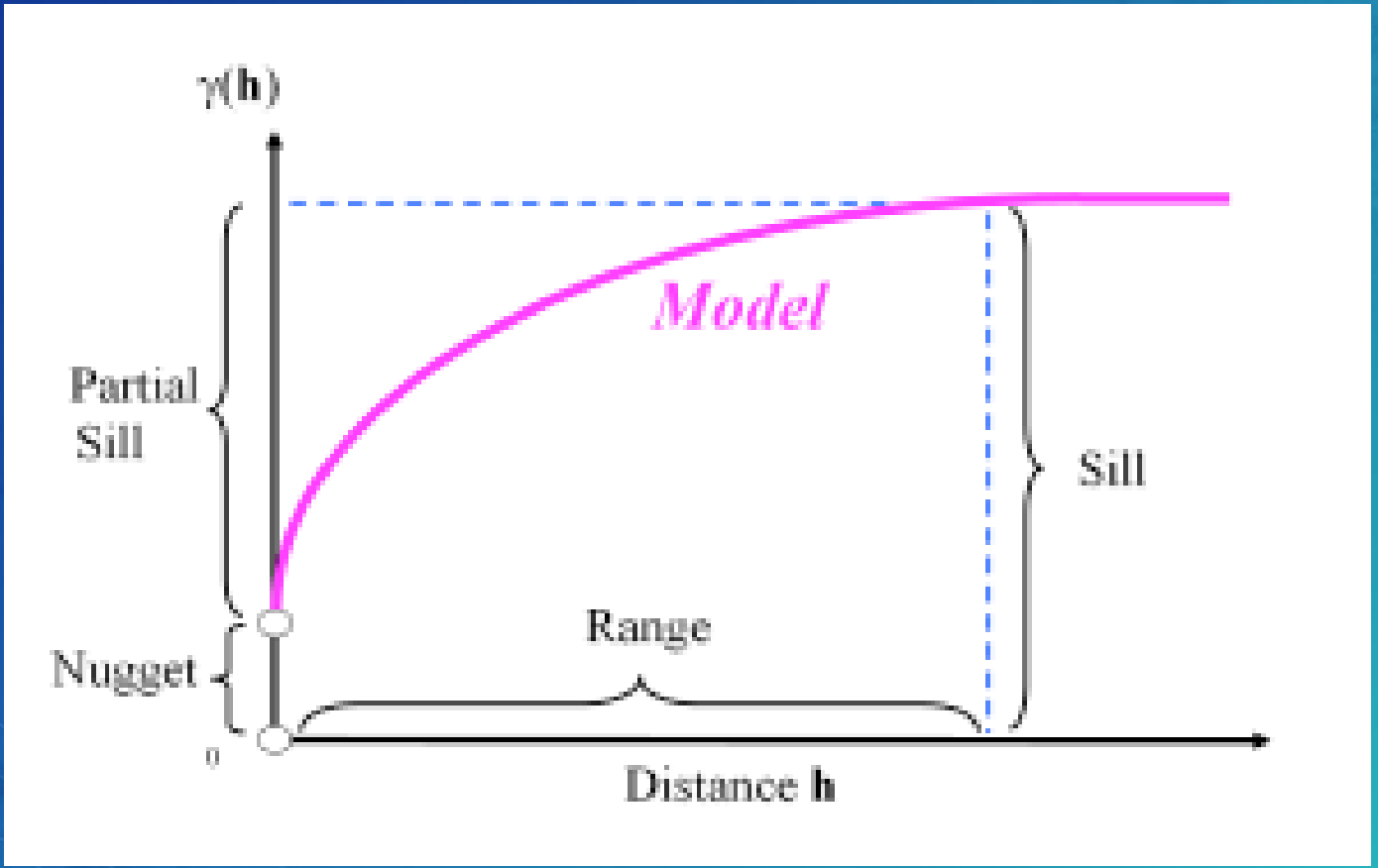
- Fits a smooth curve to the data
- Performs a quantile transformation to the normal distribution
- Performs calculations with transformed data, then transforms back at the end
- Simple kriging with normal score transformation is default kriging method



Cross-validation

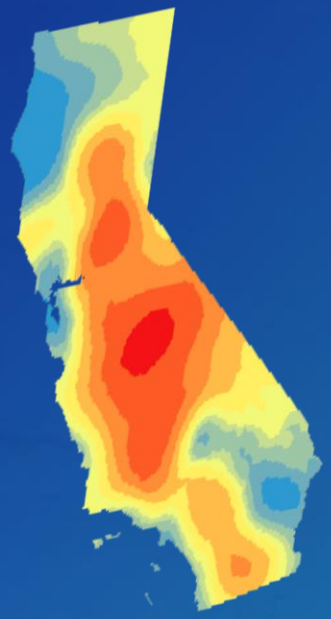
- Used to determine the reliability of the model
  - Iteratively discard each sample
  - Use remaining points to estimate value at measured location
  - Compare predicted versus measured value
- Calculates various statistics
  - Root-mean-square : root of average squared deviation from true value
    - Smaller is better
  - Mean : the average of the deviations
    - Should be close to zero
  - Root-mean-square standardized : measures whether standard errors are estimated correctly
    - Should be close to one
  - Average standard error : should be small and close to the root-mean-square

Semivariogram/Covariance Modeling

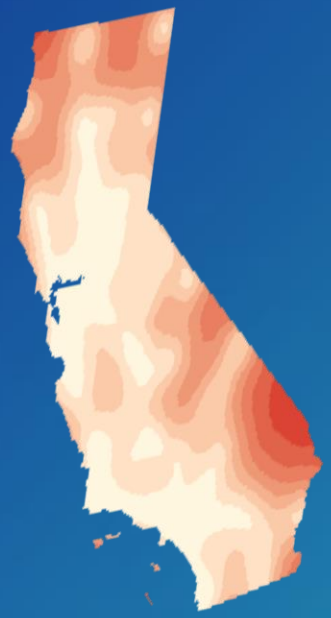


Kriging output surface types

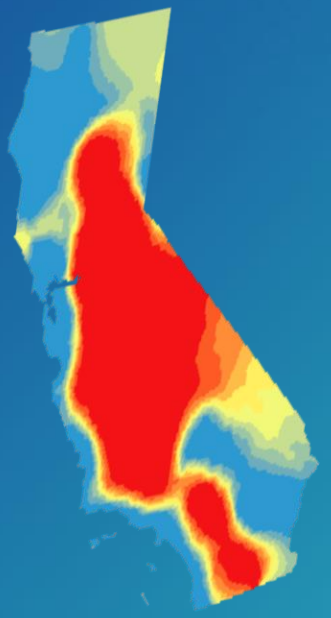
Prediction



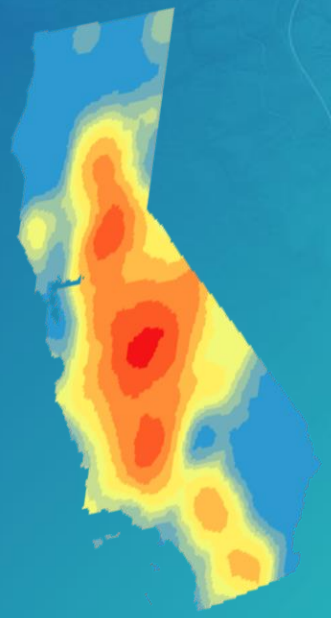
Error of Predictions

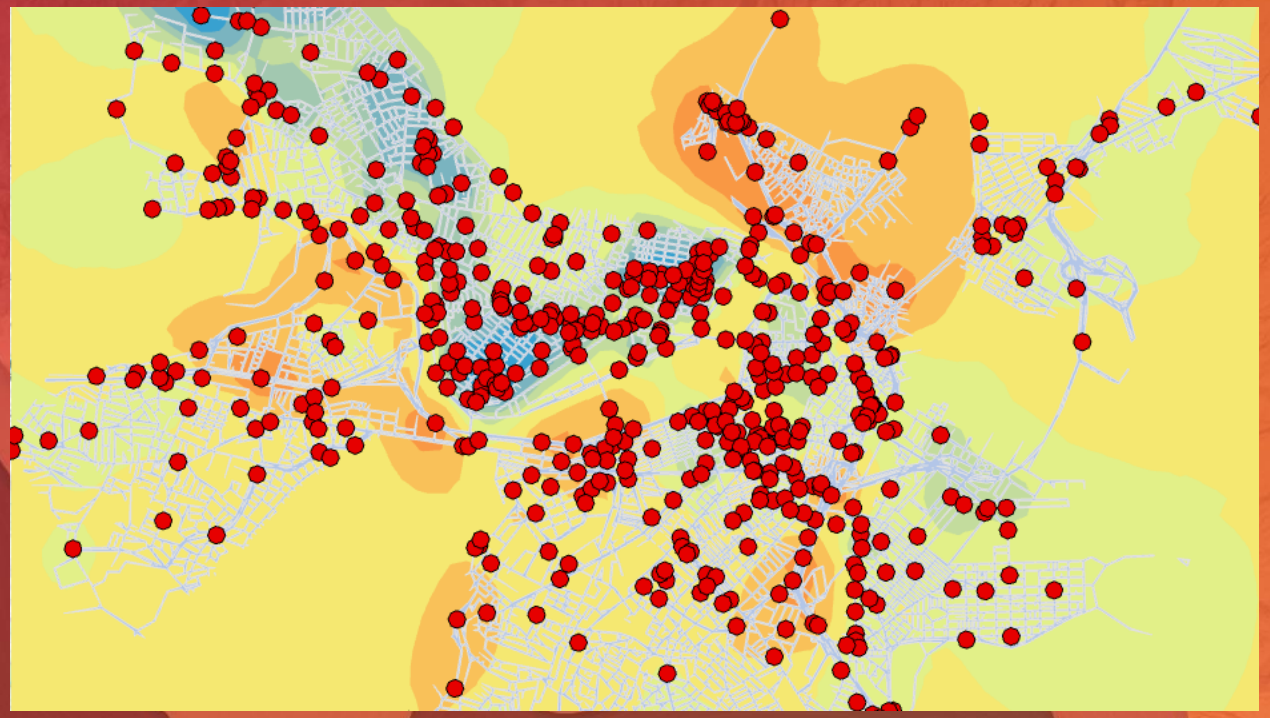


Probability



Quantile





Kriging in the Geostatistical Wizard

Demonstration

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

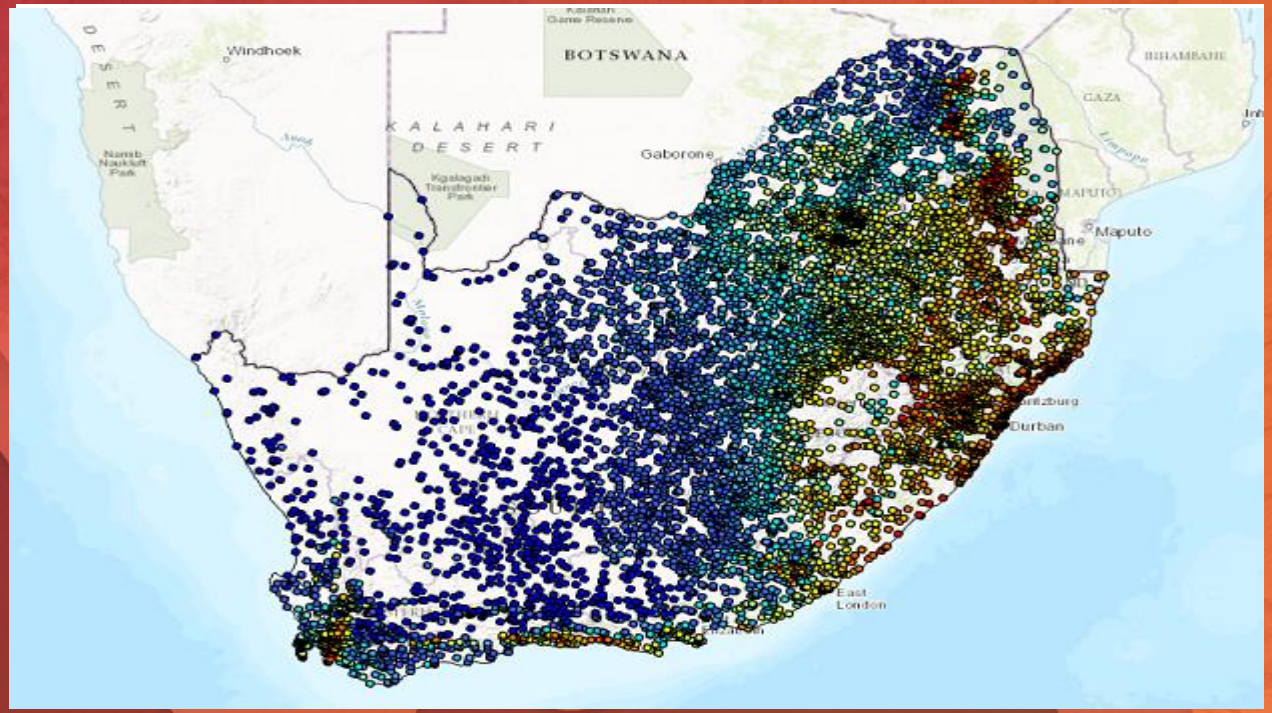
How does EBK work?

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
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

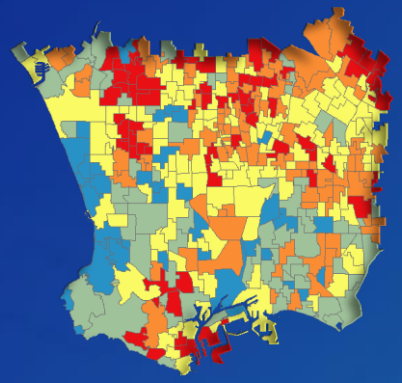
- New tool available in ArcGIS Pro 1.2
- Allows you to use explanatory variable rasters to improve predictions
- Automatically extracts useful information from explanatory variables
- Uses Principle Components to handle multicollinearity
- In ArcGIS Pro 1.3:
  - Provide your own subsets
  - New validation statistics
- In ArcGIS Pro 1.4:
  - Available in the Geostatistical Wizard



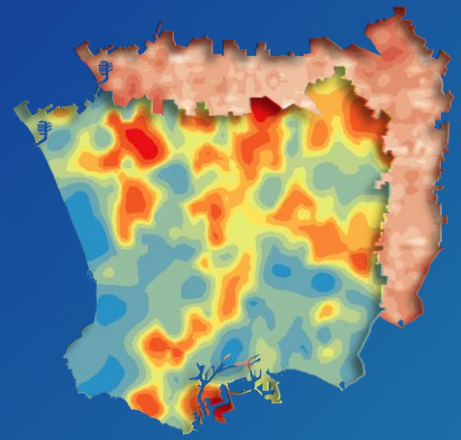
EBK and EBK Regression Prediction

Demo

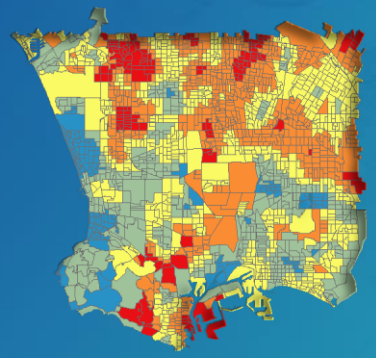
Areal Interpolation



Obesity by school zone



Obesity surface and error surface



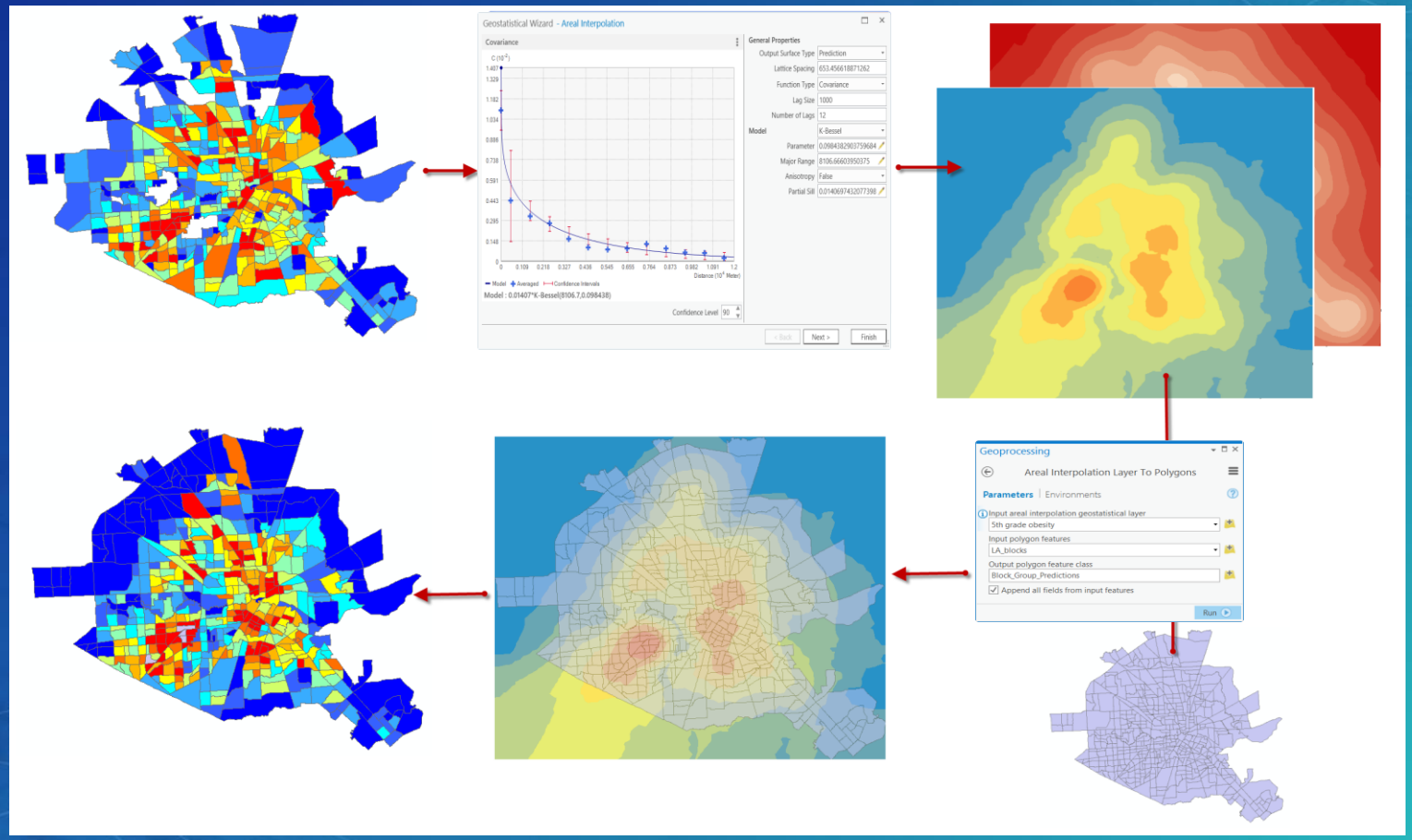
Obesity by census block

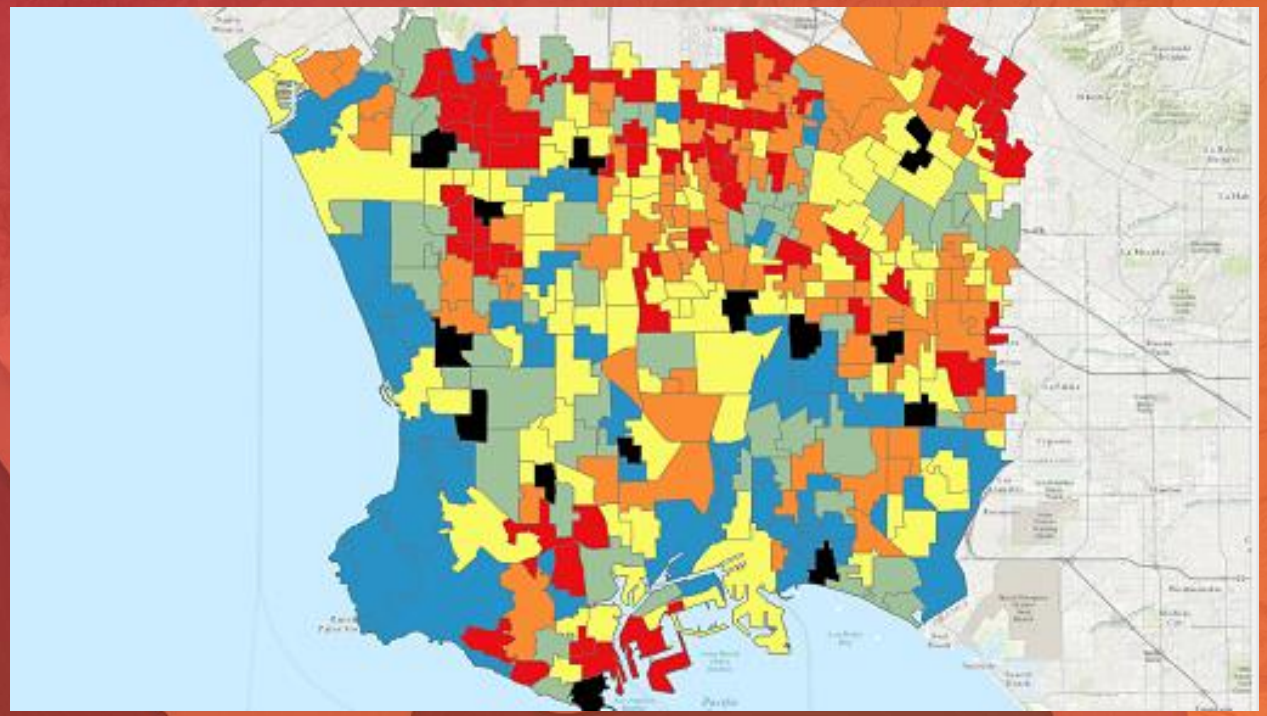
- Predict data in a different geometry
  - School zones to census tracts
- Estimate values for missing data

Types of Areal Interpolation

- Average (Gaussian)
  - Median age, average temperature
- Rate (Binomial)
  - Cancer rates, obesity rates, percent of college graduates
- Event (Overdispersed Poisson)
  - Animal counts, crimes

Areal Interpolation Workflow





Demo

Areal Interpolation


esri
THE
SCIENCE
OF
WHERE