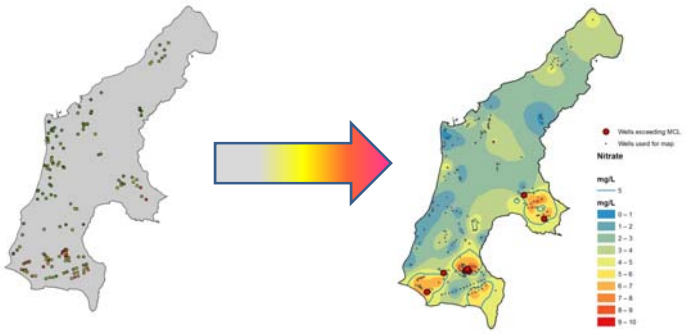
Interpolation in ArcGIS 10.1

How to predict the occurrence of values in locations where data does not exist



By: Brian G. Bearden, P.E.

For course 575.440 – GIS and Remote Sensing for Environmental Applications
Johns Hopkins University, Whiting School of Engineering
May, 2014

Presentation outline

- Interpolation methods
- How to perform interpolation in ArcGIS 10.1
- How to choose the “best” interpolation method
- How to present interpolation results

Interpolation methods

- What is interpolation?
  - The prediction of values in the spaces between data points
  - The prime focus of the field of **geostatistics** [1]
- Basic principle of geostatistics:
  - Points that are closer together are more alike than points that are farther apart [1]

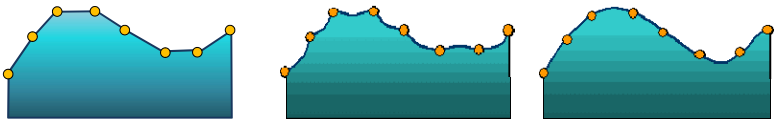
Interpolation methods

- Two categories:
  - Deterministic:
    - "exact" interpolators: surface passes through each data point
    - Good for data where points are constant, e.g. elevation, modeling outputs (such as air emissions modeling)
    - Examples: Triangulated Irregular Network (TIN), Inverse Distance Weighting (IDW), Radial Basis Functions (RBF)
  - Statistical:
    - Creates surface based on statistical analysis of surrounding data – inexact, but with quantification of error
    - Good for noisy data, e.g. environmental sampling results
    - **Kriging** is main method – several sub-varieties

Intperpolation methods

- Deterministic methods
  - Surface passes through every data point
  - Not good for noisy data

TIN | IDW | RBF



Triangulated irregular network

Inverse distance weighted

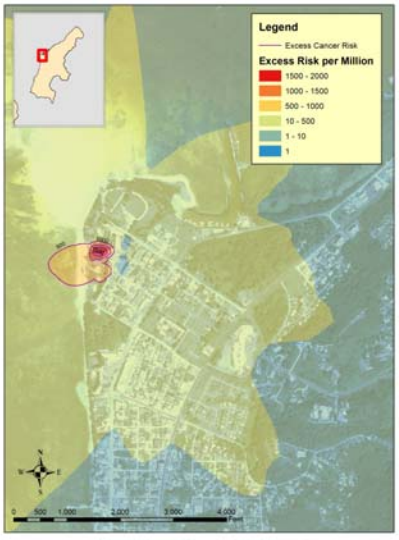
Radial basis functions (e.g., spline with tension)

Images: ESRI

Intperpolation methods

- Example:

Modeling results
  - Output from EPA air emissions model (AERMOD) in x,y,z format
  - Modeled as a TIN surface in ArcGIS
  - Good result because many datapoints, and smooth surface resulting from mathematical function



Intperpolation methods

- Probabilistic methods - Kriging
  - Statistical methods for fitting smooth surface to noisy data – good for environmental data
    - Can therefore predict data – originally developed for finding bodies of ore from isolated samples [2]

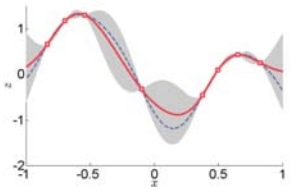


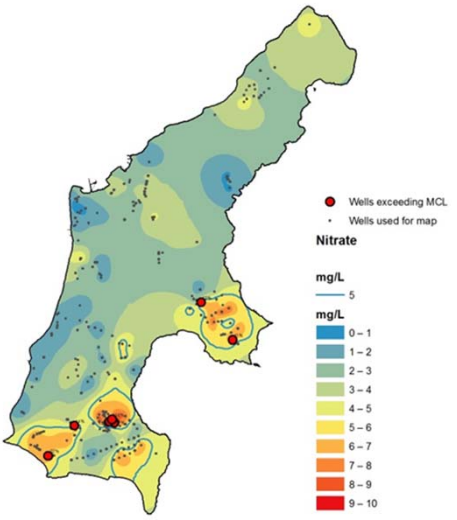
Image: Wikipedia

  - Requires data to be normally distributed – environmental data often requires transformation (e.g., log normal) [1]

Intperpolation methods

- Example:

Sampling data
  - Nitrate concentrations in wells
  - Inherently noisy data
  - Kriging interpolates a smooth surface between points AND predicts surface beyond points, based on trends suggested by surrounding data points



Interpolation methods

- Main types of Kriging:
  - Ordinary Kriging – default model in previous versions of ArcGIS
  - Simple Kriging – default model in ArcGIS 10.1 [3]
  - Universal Kriging – for anisotropic data (data with a direction component, e.g. a single contaminant plume that is elongated along groundwater flow)

Intperpolation methods

- Kriging methods available in ArcGIS 10.1
  - Through **Spatial Analyst** extension:
    - Ordinary Kriging
    - Universal Kriging
    - Spatial Analyst Kriging not recommended:
      - Implementation options very limited
        - » No ability to fine-tune
      - Produces a raster surface only
        - » Very limited analysis capability

Intperpolation methods

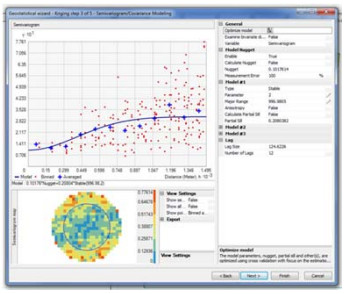
- Kriging methods available in ArcGIS 10.1
  - Through **Geostatistical analyst** extension:
    - Kriging and co-Kriking
      - Odinary
      - **Simple (Default mode – new options under 10.1)**
      - Universal
      - Indicator
      - Probability
      - Disjunctive
    - Areal interpolation (new to 10.1)
    - **Empirical Bayesian Kriging (new to 10.1)**

How to perform interpolation in ArcGIS

- Open Geostatistical Analyst toolbar
- Explore Data first
  - Check for trends. Kriging requires data to follow Gaussian (normal) distribution.
  - Environmental data often skewed – can be transformed (e.g. log tansformation)

How to perform interpolation in ArcGIS

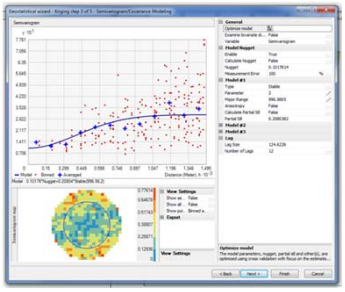
- Geostatistical Wizard to start interpolation
- Many options – for quick, reliable results choose Empirical Bayesian Kriging and log empirical transformation, if justified by the data.
- Experiment with other Kriging methods – have fun building your semivariogram models. Simple Kriging has the most options.



How to perform interpolation in ArcGIS

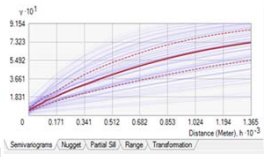
- The Kriging process:

1. Plotting the semivariogram (automatic)
2. Semivariogram modeling
3. Building prediction surface (Kriging)
4. Analyzing results



Interpolation methods

- Tip: Empirical Bayesian Kriging
  - Designed to reduce error by automating semivariogram modeling process, through statistical selection of best model from 100 randomly generated models [4],[5]



  - Easy to use, just a few decisions to make
    - Whether to transform your data (try with and without, then compare results)
    - Search neighborhood parameters – try 4-sector search, for example – can result in smoother surfaces (but longer processing time – up to 5 minutes for nitrate map with 248 sample locations)

How to choose "best" interpolation method in ArcGIS

- Right-click one of your Kriging layers and select "Compare"

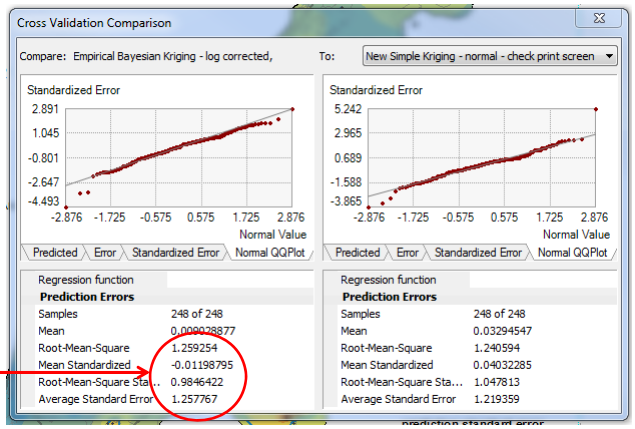
Criteria [1]:

- RMSS = 1
- Mean Stdd. = 0
- RMS = Avg. Std

Also- check transformation:

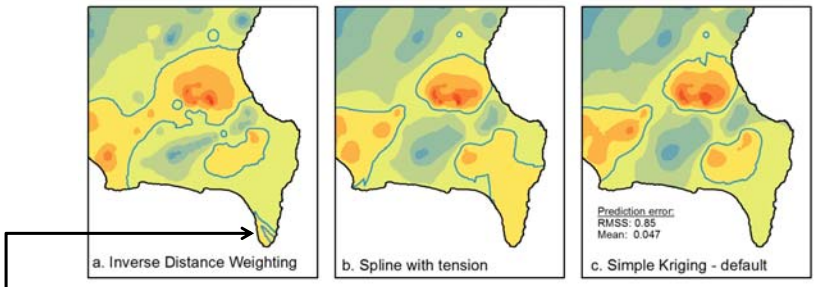
- QQPlot as near to line as possible

BEST



How to choose "best" interpolation method in ArcGIS

- Knowledge of phenomenon: does prediction make sense?

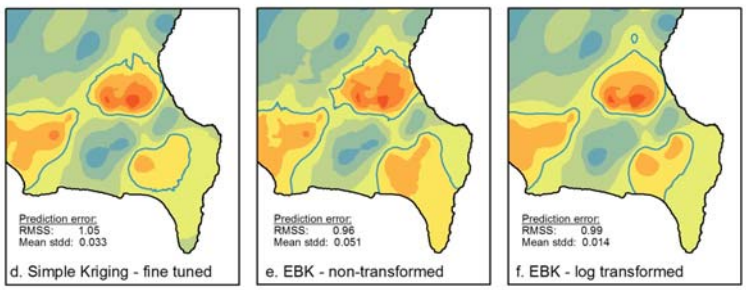


Anomalous prediction of elevated concentrations, far from nearest data point

How smooth do you think the surface should be? A noisy phenomenon, on average, should produce a relatively smooth surface.

How to choose "best" interpolation method in ArcGIS

- Ways to display model comparison results
  - On maps:



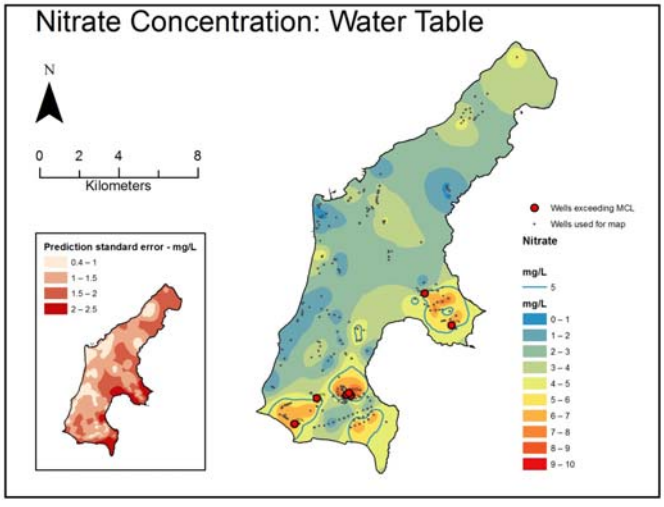
How to choose "best" interpolation method in ArcGIS

- Ways to display model validation results
  - In table:

Comparison parameters		Simple Kriging: default	Simple Kriging: fine tuned	EBK: non-transformed	EBK: log empirical transformed
Prediction errors	Target values				
Root mean square standardized (RMSS)	1	0.851	1.047	0.962	0.990
Mean standardized	0	0.0471	0.0329	0.0511	0.0136
Root mean square (RMS)	as low as possible	1.241	1.252	1.258	1.261
Average standard error	as close to RMS error as possible	1.494	1.219	1.296	1.265

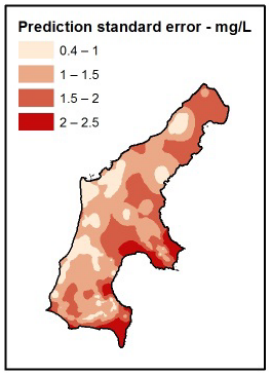
How to present interpolation results

- Sample map:



How to present interpolation results

- Prediction Standard Error
  - What is it?
    - Similar to standard deviation
    - 67% of samples collected would fall within the range of stated prediction error, +/-
    - In same units as predicted phenomenon



How to present interpolation results

- Prediction Standard Error
  - Usage and explanation [4],[6]:
    - Prediction intervals:
    - $\mu \pm z\sigma$

μ = predicted value
σ = prediction standard error
z = from table

Prediction interval	z
50%	0.67
90%	1.64
95%	1.96
99%	2.58

How to present interpolation results

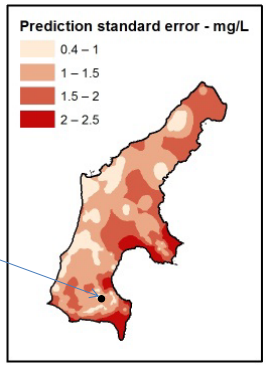
- Prediction Intervals:
  - Example (at point):
    - From prediction map,
      μ = 7.5 mg/L
    - From error map,
      σ = 1.2 mg/L
    - 95% prediction interval:

$= \mu \pm z\sigma$

$= 7.5 \pm (1.96)(1.2)$

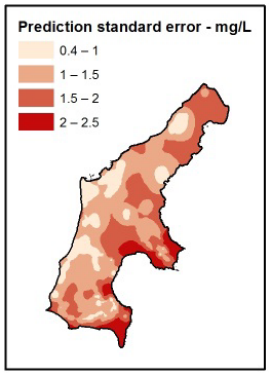
$= 7.5\, mg/L \pm 2.4\, mg/L$

$= \mathbf{5.1\ to\ 9.9\ mg/L}$



How to present interpolation results

- Prediction Standard Error
  - Good practices:
    - Use equal interval classification
    - Manually adjust to provide breaks at familiar & even values (see example to right)
    - Label legend with units



Summary

- Two main types of interpolation:
  - Deterministic (TIN, IDW, RBF)
    - Good for exact (elev.) or smooth (modeling) data
  - Probabilistic / Statistical
    - Good for noisy data (enviro. sampling)
    - Provides estimate of uncertainty
    - Useful for predicting unknown values

Summary

- Empirical Bayesian Kriging performs well & easy to use
- Comparing Kriging models:
  - Numeric (prediction error statistics)
  - Qualitative (behavior of phenomenon)
- Display & explain the prediction standard error map – always show units

References

[1] K. Johnston, J. M. Ver Hoef, K. Krivoruchko, and N. Lucas, *ArcGIS 9: Using ArcGIS Geostatistical Analyst*. Esri Press, 2003.

[2] www.kriging.com, "What is kriging anyway?" [Online]. Available: http://www.kriging.com/whatiskriging.html. [Accessed: 17-Apr-2014].

[3] Esri, "ArcGIS Help 10.1 - What's new in ArcGIS Geostatistical Analyst 10.1," 17-Dec-2012. [Online]. Available: http://resources.arcgis.com/en/help/main/10.1/index.html#//016w0000003n000000. [Accessed: 18-Apr-2014].

[4] K. Krivoruchko, "Empirical Bayesian Kriging Implemented in ArcGIS Geostatistical Analyst," 2012. [Online]. Available: http://www.esri.com/news/arcuser/1012/empirical-byesian-kriging.html. [Accessed: 25-Apr-2014].

[5] J. Pilz and G. Spöck, "Why do we need and how should we implement Bayesian kriging methods," *Stoch. Environ. Res. Risk Assess.*, vol. 22, no. 5, pp. 621–632, Aug. 2008.

[6] "Prediction interval - Wikipedia, the free encyclopedia." [Online]. Available: http://en.wikipedia.org/wiki/Prediction_interval. [Accessed: 25-Apr-2014].