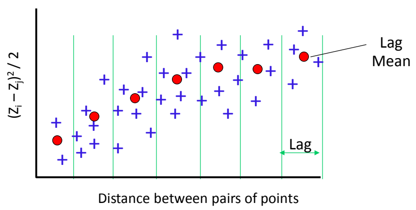
Spatial Interpolation & Geostatistics



Tobler's Law

- "All places are related, but nearby places are related more than distant places"
  - Corollary: fields vary smoothly, slowly and show strong "spatial autocorrelation" – attribute(s) and location are strongly correlated $z_i = f(x_i, y_i)$

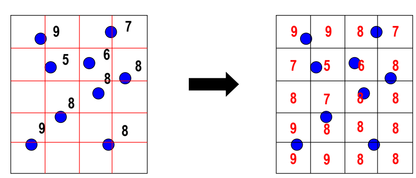
Spatial Interpolation

- Determination of unknown values or attributes on the basis of values nearby
  - Used for data that define continuous fields
    - E.g. temperature, rainfall, elevation, concentrations
    - Contouring, raster resampling are applications already discussed

Spatial Interpolation = Spatial Prediction

Spatial Interpolation

- E.g. Interpolate between variably spaced data to create uniform grid of values



Interpolation Methods

- All address the meaning of “near” in Tobler’s law differently
  - How does space make a difference?
  - Statistical mean not best predictor if Tobler’s law is true

Interpolation methods

- Inverse Distance Weighting (IDW)
  - Assumes influence of adjacent points decreases with distance

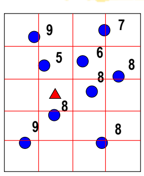
$$Z_0 = \frac{\sum_{i=1}^{n} w_i z_i}{\sum_{i=1}^{n} w_i}$$

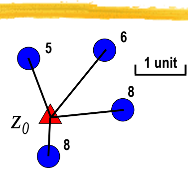
Where: Z_0 = value of estimation point

Z_i = value of neighboring point

W_i = weighting factor; e.g. = 1/(distance from neighbor)^2

Inverse Distance Weighting





On basis of four nearest neighbors:

$z_0 = (8/(1)^2 + 8/(2)^2 + 6/(2.5)^2 + 5/(2)^2)/(1.66)$

$z_0 = (8.0 + 2.0 + 0.96 + 1.25)/(1.66) = 7.36$

I.D.W.

- Unknown value is the average of the observed values, weighted by inverse of distance, squared
  - Distance to point doubles, weight decreases by factor of 4
- Can alter IDW by:
  - Alter number of closest points
  - Choose points by distance/search radius
  - Weight be directional sectors
  - Alter distance weighting; e.g. cube instead of square

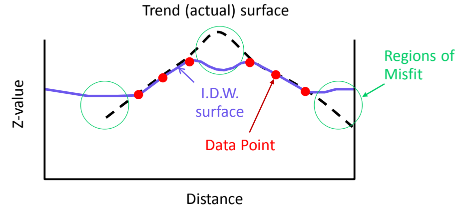
I.D.W. Characteristics

- Is an *exact method* of interpolation – will return a *measured value* when applied to measured point.
  - Will not generate smoothness or account for trends, unlike methods that are *"inexact"*
- Weights never negative –> interpolated values can never be less than smallest z or greater than largest z. "Peaks" and "pits" will never be represented.

10/28/2014 GEO327G/386G, UT Austin 9

I.D.W. Characteristics

- No peaks or pits possible; interpolated values must lie within range of known values



10/28/2014 GEO327G/386G, UT Austin 10

Interpolation Methods

- IDW is inappropriate for values that don't decrease as a function of distance (e.g. topography)
- Other deterministic techniques:
  - Spline
  - Trend

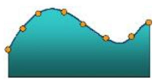
10/28/2014 GEO327G/386G, UT Austin 11

Exact Methods - Spline

- Fit *minimum curvature* surface through observation points; interpolate value from surface
- Good for gently varying surfaces
  - E.g. topography, water table heights
- Not good for fitting large changes over short distances
- Surface is allowed to exceed highest and be less than lowest measured values

10/28/2014 GEO327G/386G, UT Austin 12

Exact Methods: IDW vs. Spline



IDW: (images from ArcGIS 9.2 Help files) Spline:

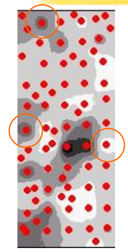
IDW:

- No predicted highs or lows above max. or min. values
- No smoothing; surface can be rough

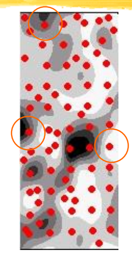
Spline:

- Minimum curvature result good for producing smooth surfaces
- Can't predict large changes over short distances

Comparisons-IDW vs. Spline



IDW, 6 nearest, contoured for 6 classes



Spline, contoured for same 6 classes

- Note smoothing of Spline – less "spikey"
- IDW contours less continuous, fewer inferred maxima and minima

Inexact (Approximate) Methods

- Trend surface –curve fitting by least squares regression
- Kriging – weight by distance, consider trends in data

Approximate Methods - Trend

- Fits a polynomial to input points using least squares regression.
- Resulting surfaces minimize variance w.r.t. input values, i.e. sum of difference between actual and estimated values for all inputs is minimized.
- Surface rarely goes through actual points
- Surface may be based on all data ("Global" fit) or small neighborhoods of data ("Local" fits).

Trend Surfaces

Equations are either:

- Linear – *1^{st} Order: fit a plane*
  - Z = a + bX + cY
- Quadratic – *2^{nd} Order: fit a plane with one bend- parabolic*
  - Z = (1^{st} Order) + dX^2 + eXY + fY^2
- Cubic – *3^{rd} Order: fit a plane with 2 bends-hyperbolic*
  - Z = (2^{nd} Order) + gX^3 + hX^2Y + iXY^2 + Y^3

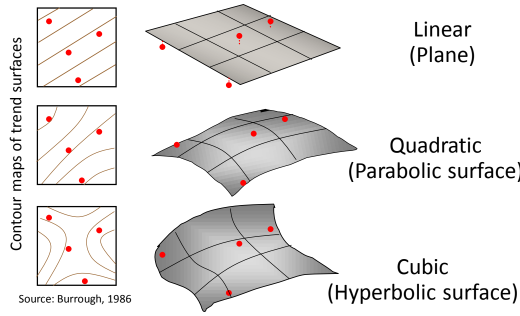
  Where:

  a, b, c, d, etc. = constants derived from solution of simultaneous equations

  X, Y = geographic coordinates

10/28/2014 GEO327G/386G, UT Austin 17

Trend Surfaces – "Global Fitting"

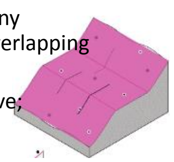


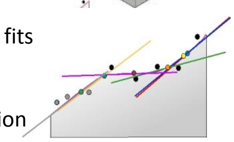
Source: Burrough, 1986

10/28/2014 GEO327G/386G, UT Austin 18

Trend Surfaces – Local fitting

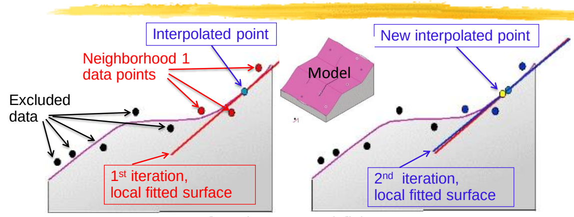
- Local Polynomial interpolation fits many polynomials, each within specified, overlapping "neighborhoods".
- Neighborhood surface fitting is iterative; final solution is based on minimizing RMS error
- Final surface is composed of best fits to all neighborhoods
- Can be accomplished with tool in ESRI Geostatistical Analyst extension





10/28/2014 GEO327G/386G, UT Austin 19

Trend Surfaces – Local fitting

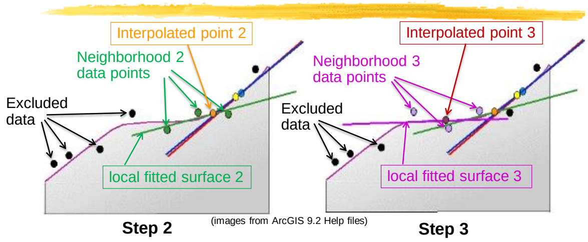


Step 1 (images from ArcGIS 9.2 Help files) Step 1; 2^{nd} Iteration

- 2-D profile view of a model surface
  - Neighborhood 1 points (red) are being fit to a plane by iteration (2 steps are shown) and an interpolated point is being created

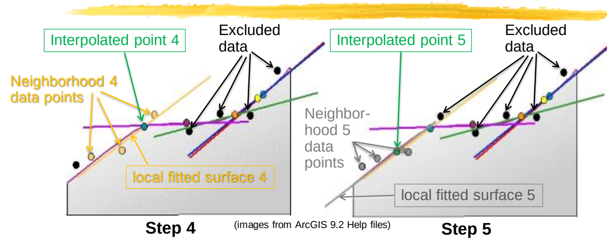
10/28/2014 GEO327G/386G, UT Austin 20

Trend Surfaces – Local fitting 2



- Model surface generated by many local fits
  - Note that several neighborhoods share some of the same data points: neighborhoods overlap

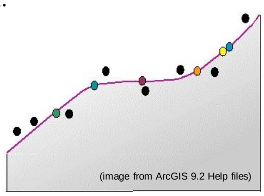
Trend Surfaces – Local fitting 3



- Five different polynomials generate five local fits; in this example all are 1^{st} Order.

Trend Surfaces – Local fitting 4

- Result:



Original data points are black

Interpolated points are in colors

- Note that model surface (purple) passes through interpolated points, not measured data points.

Why Trend, Spline or IDW Surfaces?

- No strong reason to assume that z correlated with x, y in these simple ways
- Fitted surface doesn't pass through all points in Trend
- *Data aren't used to help select model*
- → Exploratory, *deterministic* techniques, but theoretically weak

Deterministic vs. Geostatistical Models

- Deterministic: purely a function of distance
  - No associated uncertainties are used or derived
  - E.g. IDW, Trend, Spline
- Geostatistical: based on statistical properties
  - Uncertainties incorporated and provided as a result
  - Kriging

Approximate Methods - Kriging

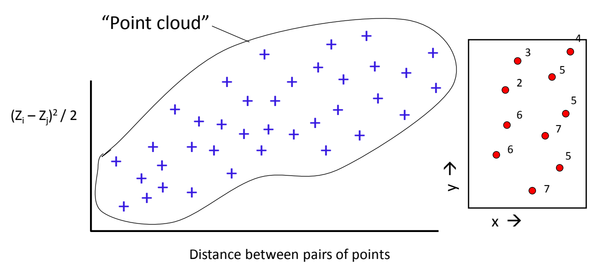
- Kriging
  - Another inverse distance method
  - Considers distance, cluster and spatial covariance (autocorrelation) – look for patterns in data
  - Fit function to selected points; look at correlation, covariance and/or other statistical parameters to arrive at weights – interactive process
  - Good for data that are spatially or directionally correlated (e.g. element concentrations)

Kriging

- Look for patterns over distances, then apply weights accordingly.
- Steps:
  1) Make a description of the spatial variation of the data - *variogram*
  2) Summarize variation by a function
  3) Use this model to determine interpolation weights

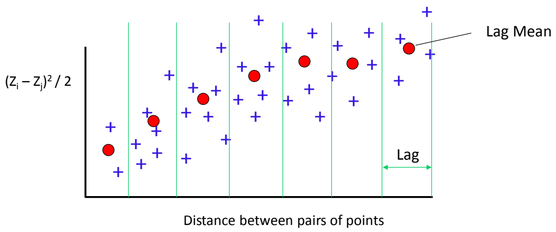
Kriging – Step 1

- Describe spatial variation with *Semivariogram*



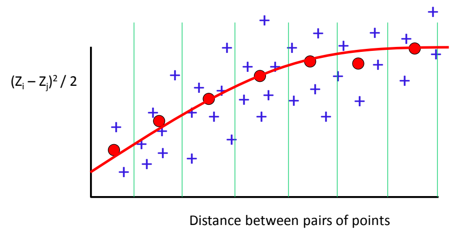
Kriging – Step 1

- Divide range into series of “lags” (“buckets”, “bins”)
- Find mean values of lags



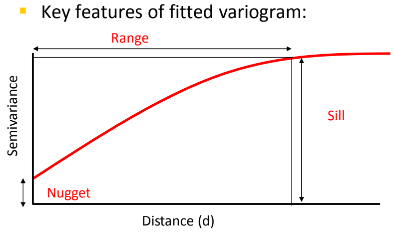
Kriging – Step 2

- Summarize spatial variation with a function
  - Several choices possible; curve fitting defines different types of Kriging (circular, spherical, exponential, gaussian, etc.)



Kriging – Step 2

- Key features of fitted variogram:



Nugget: semivariance at d = 0

Range: d at which semivariance is constant

Sill: constant semivariance beyond the range

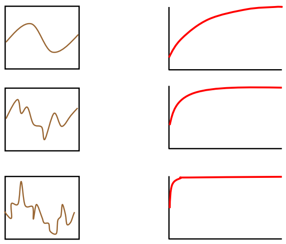
Kriging – Step 2

- Key features of fitted variogram:
  - Nugget – Measure of uncertainty of z values; precision of measurements
  - Range – No structure to data beyond the range; no correlation between distance and z beyond this value
  - Sill – Measure of the approximate total variance of z

Kriging – Step 2

- Model surface profiles and their variograms:

As local variation in surface increases, *range decreases*, *nugget increases*



Source: O'Sullivan and Unwin, 2003

Kriging – Step 3

- Determine Interpolated weights
  - Use fitted curve to arrive at weights – not explained here; see O'Sullivan and Unwin, 2003 for explanation
  - In general, nearby values are given greater weight (like IDW), but direction can be important (e.g. "shielding" can be considered)

Review: Deterministic vs. Geostatistical Models

- *Deterministic*: interpolation purely a function of distance
  - No associated uncertainties are used or derived
  - E.g. IDW, Trend, Spline
- *Geostatistical*: interpolation is statistically based
  - Uncertainties incorporated and provided as a result
    - **Kriging**

Kriging – Part II

- ***Goal:*** predict values where no data have been collected
  - Relies on first establishing:
    - DEPENDENCY – z is, in fact, correlated with distance
    - STATIONARITY – z values are stochastic (except for spatial dependency they are randomly distributed) and have no other dependence – use "detrending" or transformation tools if not Gaussian
    - DISTRIBUTION – works best if data are Gaussian. If not they have to first be made close to Gaussian.

ESRI Geostatistical Analyst Products

- Map types:
  - Prediction – contours of interpolated values
  - Prediction Standard Errors – show error distribution, as quantified by minimized RMS error (see below)
  - Probability – show where values exceed a specified threshold
  - Quantile – show where thresholds overestimate or underestimated predictions

Figure from ESRI "Intro. to Modeling Spatial Processes Using Geostatistical Analyst"

10/28/2014 GEO327G/386G, UT Austin 37

ESRI Geostatistical Analyst Products

- Maps: (e.g. max. ozone concentration, 1999)

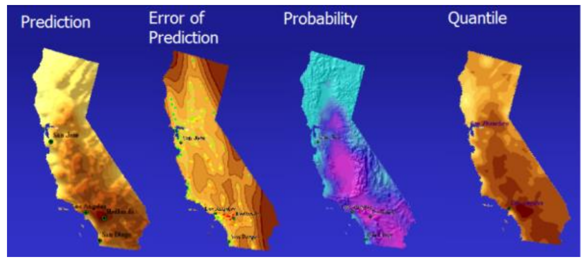
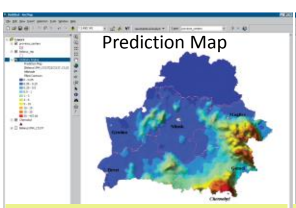


Figure from ESRI "Intro. to Modeling Spatial Processes Using Geostatistical Analyst"

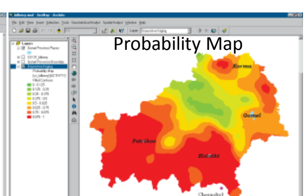
10/28/2014 GEO327G/386G, UT Austin 38

Some Kriging Products

- Prediction map – interpolated values
- Probability map- showing where critical values exceeded



Prediction map of radioceasium soil contamination levels in the country of Belarus after the Chernobyl nuclear power plant accident.



Dark orange and red indicate a probability greater than 62.5% that radiocesium contamination exceeds the critical threshold in forest berries.

Figures from ESRI "Using Geostatistical Analyst"

10/28/2014 GEO327G/386G, UT Austin 39

Kriging – Part II

- ***Goal:*** predict values where no data have been collected
  - Relies on first establishing:
    - DEPENDENCY – z is, in fact, correlated with distance
    - STATIONARITY – z values are stochastic (except for spatial dependency they are randomly distributed) and have no other dependence – use "detrending" or transformation tools if not Gaussian
    - DISTRIBUTION – works best if data are Gaussian. If not they have to first be made close to Gaussian.

10/28/2014 GEO327G/386G, UT Austin 40

1. SPATIAL DEPENDENCY

- Test with semivariogram & cross-validation plots

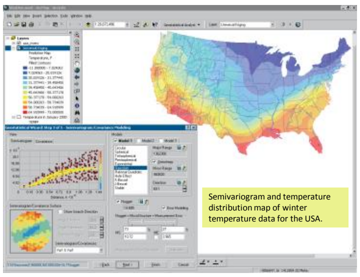
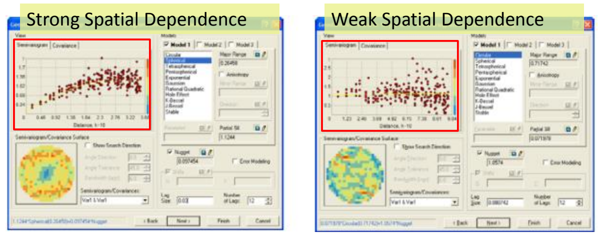


Figure from ESRI "Using Geostatistical Analyst"

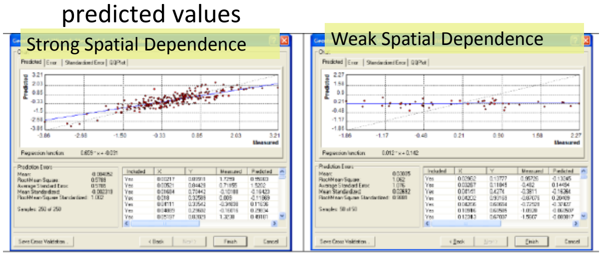
Spatial Dependence: Semivariogram



Figures from ESRI "Intro. to Modeling Spatial Processes Using Geostatistical Analyst"

Spatial Dependence: Cross-Validation Diagnostic

- Use a subset of the data to test measured vs. predicted values



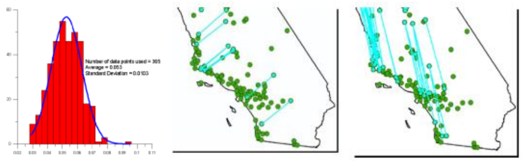
Figures from ESRI "Intro. to Modeling Spatial Processes Using Geostatistical Analyst"

Kriging – Part II

- ***Goal:*** predict values where no data have been collected
  - Relies on first establishing:
    - DEPENDENCY – z is, in fact, correlated with distance
    - STATIONARITY – z values are stochastic (except for spatial dependency they are randomly distributed) and have no other dependence – use "detrending" or transformation tools if not Gaussian
    - DISTRIBUTION – works best if data are Gaussian. If not they have to first be made close to Gaussian.

2. STATIONARITY - Randomness

- Data variance and mean is the same at all localities (or within a neighborhood of nearest points); data variance is constant in the neighborhood of investigation
- Correlation (covariance) depends only on the vector that separates localities, not exact locations, number of measurement or direction



Figures from ESRI "Intro. to Modeling Spatial Processes Using Geostatistical Analyst"

California Ozone Demo.

- Data in "Geostat_demo" folder

ArcGIS Kriging Processing Steps

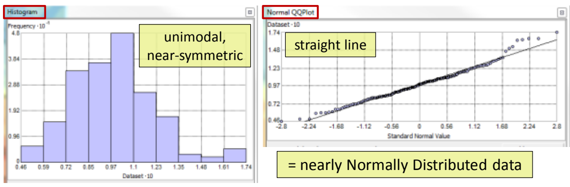
1. Add and display the data
2. Explore the data's statistical properties
3. Select a model to create a surface – make a prediction map!
4. Assess the result
5. Compare to other models

Data Exploration

1. Examine the distribution – normal (Gaussian)? Transformation to normal required?
   - *Histograms and QQPlots*
2. Identify trends, if any
   - *Trend Analysis*
3. Understand spatial autocorrelation and directional influences
   - *Semivariogram analysis*

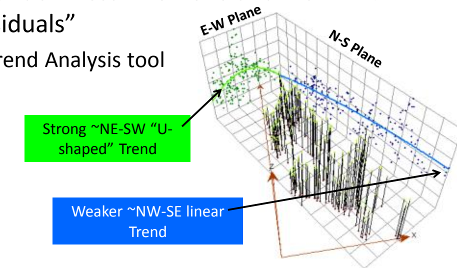
Data Exploration:
Examine the Distribution

- Normal (Gaussian) distribution? Transformation to normal required?
  - Histogram tool, QQPlot tool (compare real and standard normal distributions)



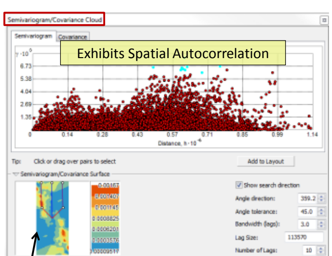
Data Exploration:
Identify Trends, If Any

- Underlying trends affect Kriging assumption of randomness – remove and work with "residuals"
  - Trend Analysis tool



Data Exploration:
Spatial Autocorrelation & Directional Influences

- Variogram Analysis:
  - Look for correlation with distance
  - Look for directional trends among pairs of points
    - Semivariogram/ Covariance Cloud tool



Directional Influence shown here – pairs in NNW-SSE direction show largest covariance over shortest distances

ArcGISKriging Processing Steps

1. Add and display the data
2. Explore the data's statistical properties
3. Select a model to create a surface – make a prediction map!
4. Assess the result
5. Compare to other models

Mapping Ozone Concentration

1. Incorporate results of Data Exploration into Model selection
   - This example:
     - remove underlying trends discovered during data exploration *that have a rational explanation.* (Analysis is then performed on residuals and trend surface is added back into final surface) = *"Detrending"*
     - Remove directional trends between pairs of points – in certain directions closer points are more alike than in other directions = *"anisotropy removal"*

Mapping Ozone Concentration – Interpolation & Cross Validation

2. Define search neighborhood for interpolation (c.f. I.D.W.)
   - Use a search ellipse (or circle) to find nearest neighbors; specify radii of ellipse, min. & max. number of points per sectors
3. Examine Cross Validation plot
   - Predicted vs. Measured for subset(s) of the data
     - "Mean error" should be close to zero
     - "RMS error" and "mean standardized error" should be small
     - "RMS standardized error" should be close to one.

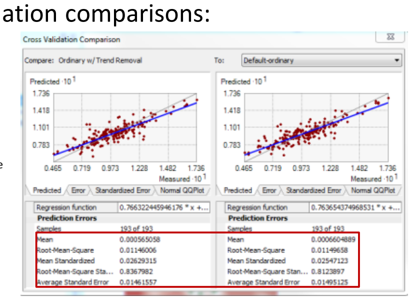
ArcGISKriging Processing Steps

1. Add and display the data
2. Explore the data's statistical properties
3. Select a model to create a surface – make a prediction map!
4. Assess the result – Cross Validation Plots
5. Compare to other models

Comparing Model Results

- Cross validation comparisons:

- "Mean error" should be close to zero
- "RMS error" and "mean standardized error" should be small
- "RMS standardized error" should be close to one.



Probality Mapping with Indicator Kriging

- Task: Make a map that show the probability of exceeding a critical threshold, e.g. 0.12 ppm ozone for an 8 hr. period
- Technique:
  - Transform data to a series of 0s and 1s according to whether they are above or below the threshold
  - Use a semivariogram on transformed data; interpret indicator prediction values as probabilities