

## GIS Workshop

### Density & Interpolation (Using Spatial Analyst & Geostatistical Analyst)

Chieko Maene  
[c-maene@northwestern.edu](mailto:c-maene@northwestern.edu)  
 Maps and State Documents Librarian  
 Government and Geographic Information and Data Services  
 University Library (Evanston)  
 Northwestern University

## Raster Analysis 2 Density Calculation (Hotspot Analysis)

### What is a Hotspot ?

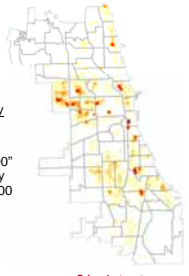
**Question: How to define a map area as 'hot' ?**

Chainey & Ratcliffe (2005) – in crime analysis,  
 “A hotspot is a geographical area of higher than average crime. It is an area of crime concentration, relative to the distribution of crime across the whole region of interest.” in *GIS and Crime Mapping*, p. 145 - 146

Popular method: “one that creates a smooth continuous surface to represent the density or volume of crimes distributed across the study area.” p. 155

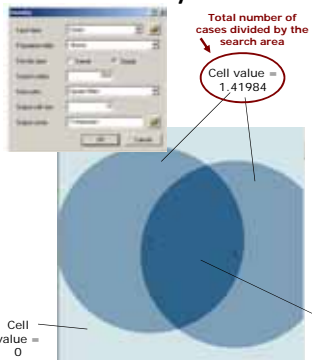
### “Simple” Hotspot (Density) Analysis

- Spatial Analyst extension is required
- It's a simple hotspot analysis because it doesn't compare the result with a "random" result...
- Use "density analysis" function
- With "Windy City" crime data
  - This data strictly for the GIS "training" purpose only
  - CPD disclaimer – not for research!
  - No quality guaranteed (i.e. human mistakes)
  - Crime locations are not accurate (i.e. "block of 1600" was geocoded as "16XX" where "XX" is a randomly assigned number, i.e. 0-99.) Error could be +/- 7,000 feet (2 blocks).
  - Data between 2/13/2005 and 2/12/2006
  - Source: Citizen Icam, <http://12.17.79.6/>



Crime hotspots

### Density Calculation - Simple



Total number of cases divided by the search area

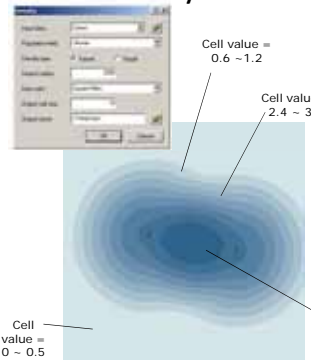
Cell value = 1.41984

- "Simple" type calculates the density of point features around each output grid cell. Units of density are points per unit of area
- The number of points that fall within the search radius area is totaled and then divided by the search radius area.
- $2,500 \text{ ft} \times 2,500 \text{ ft} \times 3.14 = 19,625,000 \text{ square ft (0.70395 square mile)}$
- $1 \text{ (# of case)} / 0.70395 \text{ sq mi (search area)} = 1.42$

Cell value = 0

Cell value = 2.83967

### Density Calculation - Kernel



Cell value = 0.6 - 1.2

Cell value = 2.4 - 3.0

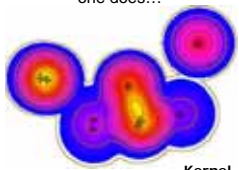
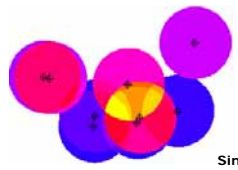
- "Kernel" type produces gradual values – or, "smoothed" values
- The surface value is highest at the location of the point, and diminishes with increasing distance
- "Kernel" calculates the density at each output cell by adding the values of all the kernel surfaces where they overlay the grid cell center. (What does that mean?)
- Thus, the output cell values cannot be explained easily!

Cell value = 4.7 ~ 5.4

Cell value = 0 ~ 0.5

### My opinion ...


- Kernel option is better for data visualization due to its smoothing effect... However, you cannot explain well what the output values mean -
- The output values make more sense with the simple method, but the output doesn't look as good as kernel one does...

Kernel Simple

### Density Calculation – Tips 1

- Output cell values will always be floating point.
- Larger values of the radius parameter produce a smoother, more generalized density raster. Smaller radius values produce a raster that shows more detail.
- Larger cell size generalize and smaller cell size show more detail (but takes more computational power + output grid would be larger..)
- If the area units are small relative to the distance between the points, the output raster values may be small. To obtain larger values, use the area unit scale factor for larger units (for example, square miles versus square feet).
- Distance unit ("not value") doesn't affect output values.

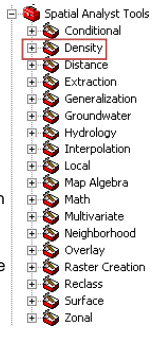


### Density Calculation – Tips 2

- Increasing the radius will not greatly change the calculated density values. (ESRI Help file) "Although more points will fall inside the larger neighborhood, this number will be divided by a larger area when calculating density. The main effect of a larger radius is that density is calculated considering a larger number of points, which can be further from the raster cell. This results in a more generalized output raster."
- My strategy
  - I prefer **Kernel** density (for better presentation.)
  - Start with relatively a larger **cell size** (not too large)
  - Find an appropriate **radius value**
    - Are there too many hotspots?? – too "tight" = not including enough neighboring cases – increase radius value
    - Are there too few hotspots?? – too generalized – decrease value
  - Decrease **cell size** for a better presentation



### ArcGIS 9 Spatial Analyst

- Key features
  - **Suitability Modeling** : Find the best location
  - **Surface Analysis** : Identify patterns or features
  - **Distance Analysis** : Calculate the cost of travel
  - **Density Analysis** : Calculate density values
  - **Hydrologic Analysis** : Derive new info with hydro & landscape data
  - **Statistical Analysis** : Model with cell calculation methods
  - **Map Algebra** : Combine multiple rasters
  - **Raster Generalization** : Reclassify or interpolate with neighboring cells



### Other considerations...

- Legend**
  - The cell value range doesn't make sense in kernel density, so it is better to simplify its legend
  - Good example! "High density" ↔ "Low density"
    - Chainey & Ratcliffe (2005)
    - [www.ojp.usdoj.gov/nij/maps/briefingbook.html](http://www.ojp.usdoj.gov/nij/maps/briefingbook.html)
- Population field**
  - Population field determines the number of times to count the point.
    - Census block center point – use "pop2000" in the population fields
    - To add a weighting factor...

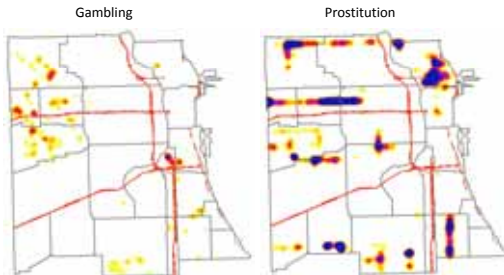



### References

- ESRI Product Overview: Spatial Analyst**  
<http://www.esri.com/software/arcgis/extensions/spatialanalyst/>  
 Includes:
  - Demos (Flash & Video)
  - Case studies
  - What's new features in ArcGIS 9 Spatial Analyst
- ArcGIS 9.3 Help**  
[http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=An overview of Spatial Analyst](http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=An%20overview%20of%20Spatial%20Analyst)
- Learning ArcGIS 9 Spatial Analyst**  
[http://training.esri.com/acb2000/showdetl.cfm?DID=6&Product\\_ID=812](http://training.esri.com/acb2000/showdetl.cfm?DID=6&Product_ID=812)  
 – First module is FREE

## Exercise 16

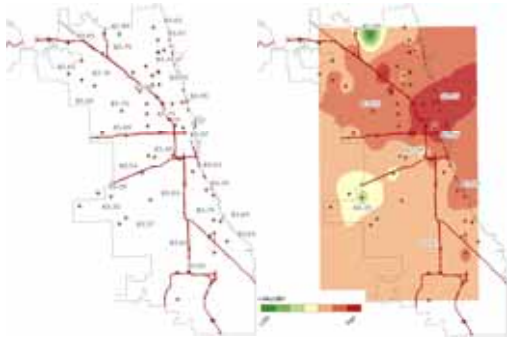
- Hot Spot (Density) Analysis



## Raster Analysis 2 Interpolation

## Interpolation

(not a good example but...)



## Interpolation

- Definition**
  - Estimation (or prediction) of surface values at unsampled points based on known surface values of surrounding points
  - A.K.A. "Surface Generalization"
- You need **sample point** data
- Applications: estimation of..
  - Elevation
  - Weather (rainfall, temperature)
  - Chemical dispersion (air pollution, ozone concentration)
  - Soil chemistry, oil and gas deposit, etc..

## Methods in Spatial Analysis

- There must be many interpolation methods developed by geographers, geologists, etc.
- ArcGIS **Spatial Analyst** used the following well-established methods only...
  - IDW (Inverse Distance Weighted)
  - Spline
  - Kriging
- Geostatistical Analyst** (although we are not going to use this extension today)
  - Gives more interpolation options (i.e. co-kriging), **plus** tools to improve value estimation.

## First Law of Geography

- The **First Law of Geography**
  - .. Is the underlying principle behind interpolation
  - "Near things are more related than distant things"
  - OR "the points near each other are more alike than those farther away – therefore, location values should be estimated based on values of points nearby"
- Also known as: **Spatial Autocorrelation**

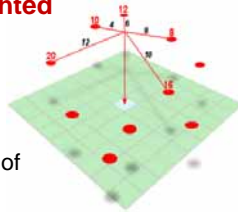
### IDW

- Inverse Distance Weighted**

- Weight will decrease as distance increases, thus “inverse”....

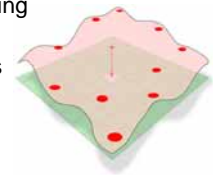
- Use averaged out weight

- The most direct application of spatial autocorrelation
- Problem: tends to create bull's eyes effect near sample points



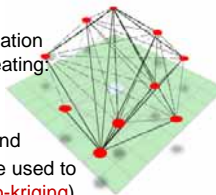
### Spline

- Has a stretching effect
- Creates a more smooth looking surface
- Sometimes estimated values exceed the values of sample data
- Two types
  - Regularized
  - Tension



### Kriging

- **Most powerful, complex and flexible of all 3 methods**
- Kriging can calculate statistical trends (deviation, fluctuation) in surface
- Kriging can assess prediction/estimation errors, and thus is also good for creating:
  - Probability surface
  - Standard error/quantile surface
- But, the method is hard to understand
- More than one layer/variable can be used to calculate more accurate surface (**co-kriging**)

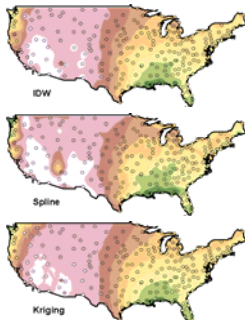


### How to determine Interpolation methods

- It is up to you...
  - You need to know your subject and surface
- Try out different methods and see which one comes closer to your ideal surface
- May depend on the quality of your sample points. If so,
  - Try to collect more sample points for areas where changes occur frequently
  - Consider Kriging (but it will require more modeling decisions...)
- **Perhaps, you need a better tool and understanding of Interpolation methods**
  - Use Geostatistical Analyst instead

### Exercise 17

- This exercise is loosely based on a free online ArcGIS lesson, **Metadata Interpolation** developed by **Carol Gersmehl**.
- Her original course material was included in the folder (note: the material was designed for the previous version of ArcView, 3.x.)
- This course incorporates her lesson concepts and data, the US weather station
- New data and instruction were also added.
- **Reference:** ESRI ArcLesson, <http://gis.esri.com/industries/education/arclessons/>

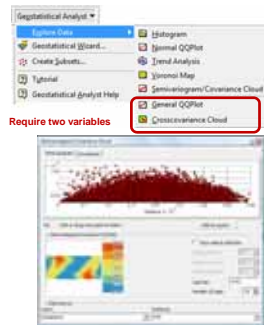


### Move on! Geostatistical Analyst

- In comparison with SA interpolation tools
  - Easy to use!
    - Interactive and visual interface (wizard)
  - Flexible!
    - More models and parameter options
  - Add confidence to your analysis
    - You will know your data and results better through step-by-step wizard tool
    - Diagnostic measurements – residual/errors, etc.

## Geostatistical Analyst

- A toolset that offers:
  - Interactive ESDA (Exploratory Spatial Data Analysis) tools
    - Learn about your data!
  - **Main tool** – Interactive modeling (surface interpolation) tools
  - Data pre-processing tools
    - I.e. data transformation
  - Model diagnostics and comparison tools



## Geostatistical Wizard

- Interactive tool that helps you to pick better methods and parameters!
- Steps:
  1. Pick a method
  2. Change parameters (power, local search radius, methods, etc.)
  3. Check errors to validate your choice of method, parameters
  4. Check the result (new raster)



## Interpolation Methods

- Deterministic interpolation
  - IDW (Inverse Distance Weighted)
  - Global polynomial interpolation
  - Local Polynomial Interpolation
  - Radial basis function
- Geostatistical interpolation
  - Kriging
    - It's statistical, so the result can tell whether or not the result is statistically significant (confidence!)
- Gaussian/statistical simulation

## Deterministic Interpolation

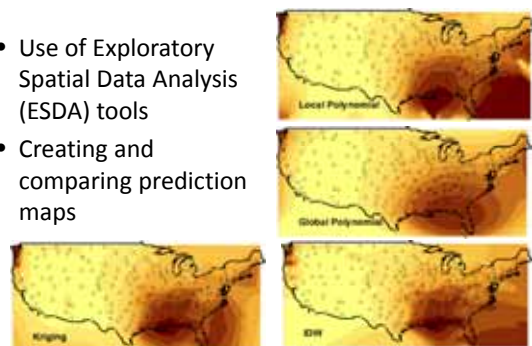
- Techniques that create predicted surfaces based on (1) the surrounding measured values (points) or (2) on specified mathematical formulas that determine the smoothness of the resulting surface
  - 1) the extent of similarity (i.e. IDW)
  - 2) the degree of smoothing (i.e. Radial Basis Functions).
- Global
  - Global techniques calculate predictions using the entire dataset. I.e. **Global Polynomial**
- Local
  - Local techniques calculate predictions from the measured points within neighborhoods, which are smaller spatial areas within the larger study area. I.e. **IDW, Local Polynomial, and Radial Basis Functions.**

## Geostatistical Interpolation

- Techniques that create predicted surfaces based on statistical models that include autocorrelation
- It will associate uncertainty with the predictions
- Pros:
  - **Not only** has the capability of producing a prediction surface, they **also** provide some measure of the certainty or accuracy of the predictions
- Cons:
  - Hard to understand and pick right parameters
  - Kriging is the only method available in ArcGIS so far

## Exercise 18

- Use of Exploratory Spatial Data Analysis (ESDA) tools
- Creating and comparing prediction maps



## References:

- I recommend scholarly books on GIS
  - My recommendation: Principles of Geographical Information Systems by Peter A. Burrough & Rachael A. McDonnell
- View demos:
  - <http://www.esri.com/software/arcgis/extensions/geostatistical/demos.html>
- Try this **free** online ESRI training course, Introduction to ArcGIS 9 Geostatistical Analyst
  - [http://training.esri.com/acb2000/showdetl.cfm?DID=6&Product\\_ID=808](http://training.esri.com/acb2000/showdetl.cfm?DID=6&Product_ID=808)
- **ArcGIS 9.3 Help**  
[http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=An overview of Geostatistical Analyst](http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=An%20overview%20of%20Geostatistical%20Analyst)