Watershed Sciences 4930 & 6920
GEOGRAPHIC INFORMATION SYSTEMS

WEEK FIVE - Lecture

VECTOR ANALYSES

Joe Wheaton

<table>
<thead>
<tr>
<th>Approximate Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIS Course(s)</td>
</tr>
<tr>
<td>4930 &amp; 6920</td>
</tr>
<tr>
<td>4930 &amp; 6920</td>
</tr>
<tr>
<td>4930 &amp; 6920</td>
</tr>
<tr>
<td>4930 &amp; 6920</td>
</tr>
<tr>
<td>4930 &amp; 6920</td>
</tr>
<tr>
<td>Spring Break – March 12 -16</td>
</tr>
<tr>
<td>4930 &amp; 6920</td>
</tr>
<tr>
<td>End of WATS 4930/6920</td>
</tr>
</tbody>
</table>

• WATS 6915... welcome to tag along for any, all or none
HOUSEKEEPING

• Forums
  - Nice job! Keep up the activity.
• Other Questions?

READING FOR TODAY

Finish reading the following by Tuesday, February 7th, 2012 (before lecture):
  • Read Chapter 9 of DeSalvo (2008) on ‘Basic Spatial Analysis’
  • Read the whole thing… pp 321-378.
TODAY’S PLAN

I. Vector Analyses
   I. Some Basics
   II. Selection
   III. Classification
   IV. Proximity Functions & Buffering
   V. Vector Overlay

SPATIAL ANALYSIS:
A way of obtaining information

- Turns raw data into useful information
  - by adding greater informative content and value
- Reveals patterns, trends, and anomalies that might otherwise be missed
- Provides a check on human intuition
  - by helping in situations where the eye might deceive
- A.K.A. Spatial Operations or Spatial Functions
  - Hundreds exist
  - All involve manipulation or calculation of coordinate or attribute values
PROJECT REQUISITE: SPATIAL ANALYSIS

• A method of analysis is **spatial** if the results depend on the locations of the objects/surface being analyzed
  - move the objects and the results change
• Method of investigation and interpretation
• Basically, anything you do with data in space to answer a question is a form of analysis
• Interpretation requires application of the right data for the task
  - how would one assess this?

SPATIAL ANALYSES CAN BE SEQUENTIAL

• The final analysis or output is arrived at after a series of spatial operations...

```
Input layer
     ↓
1st spatial operation
     ↓
1st new layer
     ↓
2nd spatial operation
     ↓
2nd new layer
     ↓
3rd spatial operation
     ↓
final layer, output
```
THREE TYPES OF OPERATIONS (VECTOR)

• Local
  - Use only data at one input location to determine value at corresponding same output location

• Neighborhood
  - Use data from both an input location plus nearby locations to determine output value

• Global
  - Use data values from entire input layer to determine each output value

1. LOCAL OPERATIONS

• Use only data at one input location to determine value at corresponding same output location
E.G. FIELD CALCULATOR - LOCAL OPERATION

- Powerful form of spatial analysis...
- You can come up with just about anything that combines existing attributes...

2. NEIGHBORHOOD OPERATIONS

- Use data from both an input location plus nearby locations to determine output value

From Bolstad (2008), Chapter 9
3. GLOBAL OPERATIONS

- Use data values from entire input layer to determine each output value

Global operation: rank order by total population in 1990

From Bolstad (2008), Chapter 9
E.G. FIELD CALCULATOR - GLOBAL OPERATION

- How would this be different than a local operation?
- E.g. sort...

NUMBER OF INPUTS & OUTPUTS VARIES

One Input - Many Outputs  |  Many Inputs - One Output

Spatial data layer 1  |  Spatial data layer 1
Function 1  |  Function 4
Spatial data layer 2  |  Spatial data layer 5
Function 2  |  Function 3
Spatial data layer 3  |  Spatial data layer 6
Spatial data layer 4  |  

- You can even have many inputs - leading to many outputs...

From Bolstad (2008), Chapter 9
TODAY’S PLAN

I. Vector Analyses
   I. Some Basics
   II. Selection
   III. Classification
   IV. Proximity Functions & Buffering
   V. Vector Overlay

VIEWS AND QUERIES

• GIS provides several different ways of viewing data:
  - maps, tables, catalog, charts

• A query is a statement or logical expression used to select elements from a larger set
  - Data Query: select records from a database
  - Spatial Query: select features by their location
• A Query is a form of selection
FOUR TYPES OF SELECTION

1. Scroll & Select
2. Set Algebra
   - Operations: less than <, greater than >, equal =, or not equal <>
3. Boolean Algebra
   - Conditions: OR, AND & NOT
4. Spatial Selection Operations
   - Adjacency & Containment

DATA QUERY AND VIEWS

You’ve already done this (perhaps unknowingly...)

States entirely north of Arkansas

States larger than 84,000 sq. km.

States both entirely north of Arkansas and larger than 84,000 sq. km.
FOUR TYPES OF SELECTION

1. Scroll & Select
2. Set Algebra
   - Operations: less than <, greater than >, equal =, or not equal <>
3. Boolean Algebra
   - Conditions: OR, AND & NOT
4. Spatial Selection
   Operations
   - Adjacency & Containment

SELECT BY ATTRIBUTES (BOOLEAN & SET ALGEBRA)

- In ArcGIS the Select by Attributes function allows us to select features based on certain values we define for a given layer using a SQL expression.

- SQL: Structured/Standard Query Language

- BASIC BOOLEAN OPERATORS:
  - ==, >, <, <>, AND (&), OR (|)

- BOOLEAN STATEMENTS always return ‘TRUE’ (1) or ‘FALSE’ (0)
**SELECT BY ATTRIBUTES** (BOOLEAN & SET ALGEBRA)

- ArcGIS Example
  - For example: Select streams with code less than or equal to 4 but not equal to 0
  - SQL: "CODE" <= 4 AND "CODE" <> 0
SELECT BY ATTRIBUTES (BOOLEAN ALGEBRA)

- ArcGIS Example
  - For example: Select streams with code less than or equal to 4 but not equal to 0
    - SQL: "CODE" <= 4 AND "CODE" <> 0

- Class Example
  - Country, State (if U.S.), Hair Color, Sex, Reside On/Off campus, Degree Seeking, Height, Age
  - “HAIR COLOR” = ‘brown’
  - “HAIR COLOR” = ‘brown’ AND age >= 21
  - We can also choose to select within selections:
    - Within the previous selection, select:
      - “STATE” <> ‘Utah’

FOUR TYPES OF SELECTION

1. Scroll & Select
2. Set Algebra
   - Operations: less than <, greater than >, equal =, or not equal <>
3. Boolean Algebra
   - Conditions: OR, AND & NOT
4. Spatial Selection Operations
   - Adjacency & Containment
EXAMPLES OF BOOLEAN EXPRESSIONS

From Bolstad (2008), Chapter 9

FOUR TYPES OF SELECTION

1. Scroll & Select
2. Set Algebra
   - Operations: less than $<$, greater than $>$, equal $=$, or not equal $<>$
3. Boolean Algebra
   - Conditions: OR, AND & NOT
4. Spatial Selection Operations
   - Adjacency & Containment

From Bolstad (2008), Chapter 9
**SPATIAL QUERIES (CONTAINMENT)**

- Features *fall inside* of some geographic area

  This feature requires special consideration. Will the query require the entire feature to fall within a given area?

**SPATIAL QUERIES (SPATIAL SELECTION)**

- Features are adjacent to one another

  Polygon A is ADJACENT to polygon B.
**SPATIAL QUERIES** (SPATIAL SELECTION)

- Features lie within a specified distance of another feature

  For example: No timber harvest within 10 meters of a stream channel

**SELECT BY LOCATION**

- With the Select By Location dialog box, you can select features based on their location relative to other features.
  - Referred to as a “Spatial Query”
  - Select by Location has several “Select From” features.
    - I want to:
      - Select features from
      - Add to the currently selected features in
      - Remove from the currently selected features in
      - Select from the currently selected features in
SELECT BY LOCATION

- Further options:
  - Intersect
  - Are within a distance of
  - Completely contain
  - Are completely within
  - ...

For example: We want to select all communication towers that lie fall within 25 Km of a BLM Wilderness Study Area.
CHARTS AND GRAPHS

Extending Selection Analyses with Charts & Graphs

• GIS Documents often require additional information to present your statement and clarify your position

• Utilizing ArcGIS tools Charts and Graphs can be inserted into ArcMap documents

• Graphs should be utilized carefully, don’t clutter your layout with unnecessary information
ARE YOU MISSING SOMETHING?

Extensions and ArcGIS

• New users often forget that there are many useful extensions that are not pre-loaded in ArcGIS

• These can be found (and activated) by right clicking on a empty space on your toolbar

• Depending on your license level, you may not have all ArcGIS extensions

• Many functions can also be found in ArcToolbox

BASIC DATA EXPLORATION

Using Geostatistical Analyst

• Statistical data exploration is required in many analysis – both basic and advanced

• Geostatistical Analyst allows a user to explore data (statistically) using basic statistical plots:
THE GEOSTATISTICAL WORKFLOW

1. Map and examine the data.
2. Pre-process data if necessary (transform, detrend, decluster).
3. Model spatial structure.
4. Define search strategy.
5. Predict values at unsampled locations.
6. Quantify uncertainty of the predictions.
7. Check that the model produces reasonable results for predictions and uncertainties.
8. Use the information in risk analysis and decision making.

DATA EXPLORATION EXAMPLES

- Histogram
- Trend Plot
- Normal Q-Q Plot
- Variogram Plot
- Semivariogram Cloud
DATA EXPLORATION EXAMPLES

NOT SURE, USE THE WIZARD
TODAY’S PLAN

I. Vector Analyses
   I. Some Basics
   II. Selection
   III. Classification
   IV. Proximity Functions & Buffering
   V. Vector Overlay

CLASSIFICATION OF SPATIAL DATA

• Classification is a spatial data operation used in conjunction with previous selection operations
• A.K.A. reclassification or recoding
• This can either create a new data set or simply a different view of same data (e.g. display properties)

From Bolstad (2008), Chapter 9
A Binary Classification

From Bolstad (2008), Chapter 9

Classification table

<table>
<thead>
<tr>
<th>state name</th>
<th>is_west</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>0</td>
</tr>
<tr>
<td>Arizona</td>
<td>1</td>
</tr>
<tr>
<td>Arkansas</td>
<td>1</td>
</tr>
<tr>
<td>Colorado</td>
<td>1</td>
</tr>
<tr>
<td>Connecticut</td>
<td>0</td>
</tr>
<tr>
<td>Wyoming</td>
<td>1</td>
</tr>
</tbody>
</table>

States west of the main branch of the Mississippi River assigned 1, east of the River assigned 0

Classifications based on distribution of data

From Bolstad (2008), Chapter 9

- Neighborhoods
  - 1074 polygons
  - population for neighborhoods ranges from 0 to 5133 (3 outliers > 3300)

Bar graph shows frequency of neighborhood population, e.g., there are 84 neighborhoods with a population between 3000 and 3100
**EQUAL-INTERVAL CLASSIFICATION**

From Bolstad (2008), Chapter 9

**NATURAL BREAKS CLASSIFICATION**

From Bolstad (2008), Chapter 9
HOW DO WE (PRI MARI LY) CLASSIFY VECTOR DATA IN ARCGIS?

EQUAL INTERVAL

Percent Population Under 5

- 3% - 6%
- 6% - 8%
- 8% - 12%
- 12% - 15%
- 15% - 18%

**DEFINED INTERVAL**

Percent Population Under 5
- 1% - 4%
- 5% - 8%
- 9% - 12%
- 13% - 16%
- 17% - 18%

**QUANTILE**

Percent Population Under 5
- 2.7% - 6.3%
- 6.3% - 6.8%
- 6.8% - 7.2%
- 7.2% - 7.8%
- 7.8% - 10%
NATURAL BREAKS

Percent Population Under 5
- 3% - 6%
- 6% - 7%
- 7% - 8%
- 8% - 10%
- 10% - 18%

GEOMETRIC INTERVAL

Percent Population Under 5
- 2.48399% - 5.1977%
- 5.1977% - 7.6279%
- 7.6279% - 9.8315%
- 9.8315% - 13.0799%
- 13.0799% - 17.9599%
STANDARD DEVIATION

Population Under 5
Deviations from Mean
- ≤ 2.5 Std. Dev.
- 2.5 - 4.5 Std. Dev.
- 4.5 - 6.5 Std. Dev.
- > 6.5 Std. Dev.

SAME DATA... JUST DIFFERENT CLASSIFICATION
TODAY’S PLAN

I. Vector Analyses
   I. Some Basics
   II. Selection
   III. Classification

IV. Proximity Functions & Buffering

V. Vector Overlay
PROXIMITY & BUFFER

• Proximity functions modify existing features or create new features that depend on distances from that feature in some way....

BUFFERS

• Create a new object consisting of areas within a user-defined distance of an existing object (point, line, or polygon)
  - e.g., to determine areas impacted by a proposed highway
  - e.g., to determine the service area of a proposed hospital
• Feasible using either rasters or vectors
**SPATIAL QUERIES** (SPATIAL SELECTION)

- YOU’VE SEEN THIS... Features lie within a specified distance of another feature

  For example: No timber harvest within 10 meters of a stream channel

**TYPES OF SIMPLE VECTOR BUFFERS....**

- ‘Simple’ refers to fixed distance...

  - **points**
  - **lines**
  - **polygons**

*From Bolstad (2008), Chapter 9*
ALLOWS THEN SELECTION QUERIES OF:

- Enclosed areas, inside buffer areas, outside buffer areas...

FANCIER... VARIABLE DISTANCE BUFFER

<table>
<thead>
<tr>
<th>river_identifier</th>
<th>buffdist</th>
</tr>
</thead>
<tbody>
<tr>
<td>mississippi</td>
<td>100</td>
</tr>
<tr>
<td>missouri</td>
<td>50</td>
</tr>
<tr>
<td>arkansas</td>
<td>50</td>
</tr>
<tr>
<td>ohio</td>
<td>75</td>
</tr>
<tr>
<td>tennessee</td>
<td>75</td>
</tr>
<tr>
<td>st. croix</td>
<td>75</td>
</tr>
<tr>
<td>illinois</td>
<td>75</td>
</tr>
<tr>
<td>wisconsin</td>
<td>75</td>
</tr>
</tbody>
</table>
BUFFER ISSUES TO CONSIDER

• If you have many features to buffer, what happens if buffer zones overlap?

• If lines, are ends rounded or square?

• If polygons, outside or inside?

• If outside, do you want to include the original features?

TODAY’S PLAN

I. Databases
II. Vector Analyses
   I. Some Basics
   II. Selection
   III. Classification
   IV. Proximity Functions & Buffering
   V. Vector Overlay
SIMPLY Overlaying DIFFERENT Types...

- Five possible combinations
- Sometimes just simple combination
- Output is always the simpler data type

From Bolstad (2008), Chapter 9

SPECIAL CASES OF POLYGON OVERLAYS

From Bolstad (2008), Chapter 9
VECTOR OVERLAYS IN ARC

- ArcToolbox → Analysis Tools → Extract|Overlay
- ArcToolbox → CoverageTools → Analysis → Extract|Overlay
- Extract: SELECT, CLIP, SPLIT
- Overlay: ERASE, IDENTITY, INTERSECT, SYMMETRICAL DIFFERENCE, UNION, UPDATE

VECTOR OVERLAYS: DISCRETE OBJECTS

In this example, two polygons are overlaid to form 9 new polygons:

- Both input polygons (1) Polygon A only (4) Polygon B only (4)

Most overlay operations differ in two respects:
1. Which of these polygons are kept or discarded?
2. What happens to the attributes?
VECTOR OVERLAYS: DISCRETE OBJECTS

A CLIP operation would keep the features and attributes of A that fell within B

An INTERSECT would keep the overlapping areas and retain attributes from both

Clip: A and B
Intersect: A and B + attributes

Erase: A only
PROBLEMS WITH FIELD OVERLAYS

- In any two such layers there will almost certainly be boundaries that are common to both layers
  - e.g. following rivers, coastlines, county lines
- The two versions of such boundaries will not be coincident
- As a result large numbers of small sliver polygons will be created
  - these must somehow be removed
  - this is normally done using a user-defined tolerance
- Rasters do not have this problem
FOR EXAMPLE, SLIVERS...

common boundary

slivers

TODAY’S PLAN

I. Databases
II. Vector Analyses
   I. Some Basics
   II. Selection
   III. Classification
   IV. Proximity Functions & Buffering
   V. Vector Overlay
THIS WEEK’S LAB

VECTOR ANALYSES

• Another Oil Spill!
• The city needs your GIS skills to help respond to the disaster