Watershed Sciences 4930 & 6920
ADVANCED GIS

VECTOR ANALYSES

Joe Wheaton
HOUSEKEEPING

- No In-Class Lecture This Week
- What’s this?
**WATS 4930/6920... WHERE WE’RE GOING**

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

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>Dates</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Vector Analyses</td>
<td>Feb 4 &amp; 6</td>
<td>5. Vector Analysis</td>
</tr>
<tr>
<td>6</td>
<td>Raster Analyses</td>
<td>Feb 13</td>
<td>6. Working w/ DEMs</td>
</tr>
<tr>
<td>7</td>
<td>Raster Analyses</td>
<td>Feb 20</td>
<td>7. Building DEMs</td>
</tr>
<tr>
<td>8</td>
<td>GIS Modeling</td>
<td>Feb 27</td>
<td>8. Morphometric Analyses</td>
</tr>
<tr>
<td>9</td>
<td>Collecting Your Own Data</td>
<td>Mar 4 &amp; 6</td>
<td>9. Blimp &amp; Georeferencing LabLab</td>
</tr>
<tr>
<td></td>
<td>Spring Break – March 10-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>GIS Modeling &amp; Synthesis</td>
<td>Podcast</td>
<td>10. Habitat Modeling</td>
</tr>
<tr>
<td>11</td>
<td>Project Proposals</td>
<td>Mar 27</td>
<td></td>
</tr>
</tbody>
</table>

End of WATS 4930/6920
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
  - Thousands exist
  - All involve manipulation or calculation of coordinate or attribute values
MORE INFORMATION?

BLOOMS TAXONOMY

- **Evaluation**: Assessing theories; Comparison of ideas; Evaluating outcomes; Solving; Judging; Recommending; Rating
- **Synthesis**: Identifying and analyzing patterns; Organisation of ideas; recognizing trends
- **Analysis**: Understanding; Translating; Summarising; Demonstrating; Discussing
- **Application**: Using and applying knowledge; Using problem solving methods; Manipulating; Designing; Experimenting
- **Comprehension**: Recall of information; Discovery; Observation; Listing; Locating; Naming
- **Knowledge**: Using old concepts to create new ideas; Design and Invention; Composing; Imagining; Inferring; Modifying; Predicting; Combining
SPATIAL ANALYSES CAN BE SEQUENTIAL

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

From Bolstad (2008), Chapter 9
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

From Bolstad (2008), Chapter 9
1. LOCAL OPERATIONS

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

From Bolstad (2008), Chapter 9
• 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

Neighborhood operation: number of adjacent states
FOR EXAMPLE... NEAR- NEIGHBORHOOD OPS.
3. GLOBAL OPERATIONS

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

Global operation: rank order by total population in 1990
• How would this be different then a local operation?
• E.g. sort...
NUMBER OF INPUTS & OUTPUTS VARIES

One Input - Many Outputs

- Spatial data layer 1
  - function 1
    - Spatial data layer 2
    - function 2
      - Spatial data layer 3
      - function 3
        - Spatial data layer 4

Many Inputs - One Output

- Spatial data layer 1
  - function 3
    - Spatial data layer 5
  - function 4
    - Spatial data layer 6

- 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
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...)

http://resources.arcgis.com/en/help/main/10.2/index.html#/Selecting_features_interactively/00s50000000w000000
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
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)

From Bolstad (2008), Chapter 9
SELECT BY ATTRIBUTES

SELECT * FROM SGID500_Streams WHERE

"CODE" < 4 AND "CODE" <> 0
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
• 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
• 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.
SELECT BY LOCATION

Select By Location

Let you select features from one or more layers based on where they are located in relation to the features in another layer.

I want to:

select features from

the following layer(s):

- [ ] SGID100_CommunicationTowersGNIS
- [X] SGIDMISC_CacheCoMonuments
- [ ] SGID500_Streams

Only show selectable layers in this list

that:

are within a distance of

the features in this layer:

SGID024_BLMWildernessStudyAreas

Use selected features

[ ] Apply a buffer to the features in SGID024_BLMWildernessStudyAreas

of 25 Kilometers

Preview:
The red features represent the features in SGID024_BLMWildernessStudyAreas. The highlighted cyan features are selected because they are within a distance of the red features.

Points

Lines

Polygons

Apply

Close
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

http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Hiding_and_showing_toolbars/00v00000001000000/
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
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 the same data (e.g., display properties).

From Bolstad (2008), Chapter 9.
A BINARY CLASSIFICATION

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

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>....</td>
<td>...</td>
</tr>
<tr>
<td>Wyoming</td>
<td>1</td>
</tr>
</tbody>
</table>
CLASSIFICATIONS BASED ON DISTRIBUTION OF DATA

- 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

- 0 - 1130
- 1130 - 2156
- 2156 - 5133

Histogram showing frequency of population.
HOW DO WE (PRIMARILY) CLASSIFY VECTOR DATA IN ARCGIS?
EQUAL INTERVAL

Percent Population Under 5

- 3% - 6%
- 6% - 9%
- 9% - 12%
- 12% - 15%
- 15% - 18%

DEFINED INTERVAL

Percent Population Under 5

- 1% - 4%
- 5% - 8%
- 9% - 12%
- 13% - 16%
- 17% - 20%

Classification
- Method: Defined Interval
- Interval Size: 400

Classification Statistics
- Count: 254
- Minimum: 0.08716
- Maximum: 2673
- Sum: 25500
- Mean: 100.4
- Median: 21.28
- Standard Deviation: 293.5

Break Values

QUANTILE

Percent Population Under 5

- 2.7% - 6.3%
- 6.3% - 6.8%
- 6.8% - 7.2%
- 7.2% - 7.8%
- 7.8% - 18%
NATURAL BREAKS

Percent Population Under 5

- 3% - 6%
- 6% - 7%
- 7% - 8%
- 8% - 10%
- 10% - 18%
GEOMETRIC INTERVAL

Percent Population
Under 5

- 2.6539% - 5.1977%
- 5.1978% - 7.0378%
- 7.0379% - 9.5815%
- 9.5816% - 13.0979%
- 13.098% - 17.9589%

Population Under 5 Deviation from Mean

- < -2.5 Std. Dev.
- -2.5 - -1.5 Std. Dev.
- -1.5 - -0.5 Std. Dev.
- -0.5 - 0.5 Std. Dev.
- 0.5 - 1.5 Std. Dev.
- 1.5 - 2.5 Std. Dev.
- > 2.5 Std. Dev.

Specify the interval size as a fraction of standard deviations. Smaller fractions generate more classes.
SAME DATA... JUST DIFFERENT CLASSIFICATION
HISTOGRAMS! & OTHERS…
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....

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

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

Line features

Buffer

outside buffer

enclosed area

inside buffer

From Bolstad (2008), Chapter 9
FANCIER... VARIABLE DISTANCE BUFFER

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

<table>
<thead>
<tr>
<th>Input layer 1</th>
<th>Input layer 2</th>
<th>Output layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>point</td>
<td>line</td>
<td>point</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1-</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>2B</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3-</td>
</tr>
<tr>
<td>point</td>
<td>polygon</td>
<td>point</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1A</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2A</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3B</td>
</tr>
<tr>
<td>line</td>
<td>line</td>
<td>line</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>line</td>
<td>polygon</td>
<td>line</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>10R</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>11S</td>
</tr>
<tr>
<td>polygon</td>
<td>polygon</td>
<td>polygon</td>
</tr>
<tr>
<td>100</td>
<td>200</td>
<td>100R</td>
</tr>
<tr>
<td>100</td>
<td>200</td>
<td>100S</td>
</tr>
<tr>
<td>100</td>
<td>200</td>
<td>200S</td>
</tr>
</tbody>
</table>

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 → Coverage Tools → Analysis → Extract|Overlay
- Extract: SELECT, CLIP, SPLIT
- Overlay: ERASE, IDENTITY, INTERSECT, SYMMETRICAL DIFFERENCE, UNION, UPDATE
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?
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
Union:
A only | B only | A and B + attributes
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
Finish reading the following by Thursday, February 7th, 2013 (before lecture):

- Read Chapter 9 of Bolstad (2008) on 'Basic Spatial Analysis'
- Read the whole thing... pp 321-378.
THIS WEEK’S LAB

VECTOR ANALYSES

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