Spatial Analysis

A. Michelle Lawing
Ecosystem Science and Management
Texas A&M University
College Station, TX 77843
alawing@tamu.edu
A picture is worth a thousand words

Diagram:
- Input layer
- 1st spatial operation
  - 1st new layer
  - 2nd spatial operation
- 2nd new layer
  - 3rd spatial operation
- Final layer, output
Draw a diagram

• Take out a blank sheet of paper

• Draw or diagram the steps you took to get to school today

• Have at least 5 steps, such as, turn off alarm, get out of bed, get dressed, etc...
I/O – input/output

One Input - Many Outputs

Spatial data layer 1

- function 1
- function 2

Spatial data layer 2

- function 2

Spatial data layer 3

- function 3

Spatial data layer 4

Many Inputs - One Output

Spatial data layer 1

Spatial data layer 5

Spatial data layer 6

function 4
Input scope

Local operation: 1990 population density

Neighborhood operation: number of adjacent states

Global operation: rank order by total population in 1990
What is spatial analysis?

The process of identifying spatial patterns in data and spatial relationships between features to understand phenomena.
Taxonomy of data types

• Point patterns – discrete points in space (e.g., an event represented by x,y)

• Continuous surfaces – a blanket of data (e.g., a gradient) that can be regularly or irregularly distributed

• Aggregates – represents summaries of discrete or continuous data within an area or polygon
Modeling process

• Exploratory analysis
  – Visualization of data
  – Summarizing data
  – Identifying spatial dependency in data

• Inferential models
  – Testable hypotheses
  – Based on nature of data
Point pattern analysis

- The location of the points is the object of interest
Surface Analysis

- The objective is to reconstruct the surface from which the samples were removed and measured.
Areal analysis

- E.g., census data
- Delimited by closed polygons
- Assumes internal homogeneity
Three “laws” of geographic information

• Spatial dependency

• Spatial heterogeneity

• Spatial problems are scale dependent
Spatial dependency

• First law of geography, “everything is related to everything else, but near things are more related than distant things” Waldo Tobler – emeritus Prof at UCSB

• Spatial autocorrelation
  – Moran’s I
  – Spatial Variogram
Spatial autocorrelation

County-based percentages of groundwater NO$_3$ observations exceeding NO$_3$ maximum contaminant level as indicated by Moran’s I and LISA (Local Indicators of Spatial Autocorrelation)
Spatial regression

- Capture spatial dependency in regression analysis
Spatial interpolation

- Estimate variables at unobserved locations based on values at observed locations
- Spatial regression
- Inverse weighting distance
- Kriging
Spatial interpolation

Generating Temperature maps from climate stations

Necessary steps:
- overlay stations with DEM
- calculate regressions
- interpolate regionally adjusted intercepts (Intercept +/- residuals)
- use fine-scale DEM, intercept and Lpsrt to derive new climate map.
Spatial interaction

- Estimates the flow of people, material, or information between geographic localities

The geography of talk in Great Britain (total talk time)
Simulation and Modeling

- Bottom-up emergence of complex patterns starting with interactions on the individual level

http://jasss.soc.surrey.ac.uk/14/3/7.html
Operations

- Selection
- Reclassification
- Dissolving
- Buffering
- Overlay
Operations

• Selection
• Reclassification
• Dissolving
• Buffering
• Overlay
Spatial Selection

• Identifying features based on spatial criteria adjacency, connectivity, containment, or arrangement
Spatial Selection

- Adjacency depends on algorithm used

Adjacency - shared line required

Adjacency - shared node or line required
Spatial Selection

- Identifying features based on spatial criteria: adjacency, connectivity, containment, or arrangement
Spatial Selection

- Identifying features based on spatial and non-spatial criteria

Maps:
- Arkansas
- States larger than 84,000 sq. km.
- States entirely north of Arkansas
- States both entirely north of Arkansas and larger than 84,000 sq. km.
Operations

• Selection
• Reclassification
• Dissolving
• Buffering
• Overlay
Reclassification

- An assignment of a class or value based on attributes or geography of an object (e.g., parcels reclassified by size)
Reclassification

Source layer

Classification table

<table>
<thead>
<tr>
<th>In</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
</tr>
<tr>
<td>I</td>
<td>3</td>
</tr>
</tbody>
</table>

Output layer
Reclassification

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>
Reclassification

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
Reclassification
Reclassification

Equal-area classification

<table>
<thead>
<tr>
<th>Class</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 902</td>
<td>4</td>
</tr>
<tr>
<td>903 - 1223</td>
<td>6</td>
</tr>
<tr>
<td>1224 - 5133</td>
<td>3</td>
</tr>
</tbody>
</table>
Reclassification

Natural breaks classification

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

Frequency (%)

Population

0 1000 2000 3000 5000
Operations

• Selection
• Reclassification
• **Dissolving**
• Buffering
• Overlay
Dissolving

- Combining like features in a data layer

Dissolve table

<table>
<thead>
<tr>
<th>state name</th>
<th>is_west</th>
<th>dissolve value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>0</td>
<td>E</td>
</tr>
<tr>
<td>Arizona</td>
<td>1</td>
<td>W</td>
</tr>
<tr>
<td>Arkansas</td>
<td>1</td>
<td>W</td>
</tr>
<tr>
<td>Colorado</td>
<td>1</td>
<td>W</td>
</tr>
<tr>
<td>Connecticut</td>
<td>0</td>
<td>E</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Wyoming</td>
<td>1</td>
<td>W</td>
</tr>
</tbody>
</table>
Operations

- Selection
- Reclassification
- Dissolving
- Buffering
- Overlay
Buffering
Buffering

Vector buffer

Raster buffer

buffer distance

input feature

output buffer

buffer distance

input feature

output buffer
Raster buffer is array of distances

\[ \text{distance} = \sqrt{x^2 + y^2} \]

Buffer distance = 15 units

outside buffer distance
within buffer distance
target cell - 0 distance

reclassed as out
reclassed as in
reclassed as in
Vector buffers
Vector buffers

\[ r = \sqrt{(x-x_1)^2 + (y-y_1)^2} \]
Vector buffers
Operations

• Selection
• Reclassification
• Dissolving
• Buffering
• Overlay
Overlay

- Combining data layers with features and attributes

Diagram showing overlay of two layers with attributes for each layer and combined attributes for layers A & B.
Raster Overlay

**Layer A**

<table>
<thead>
<tr>
<th>Geographic data</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>A</td>
</tr>
</tbody>
</table>

**Attribute data**

<table>
<thead>
<tr>
<th>Type</th>
<th>soil_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Evard loam</td>
</tr>
<tr>
<td>B</td>
<td>Cecil clay</td>
</tr>
</tbody>
</table>

**Layer B**

<table>
<thead>
<tr>
<th>Geographic data</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

**Attribute data**

<table>
<thead>
<tr>
<th>ID</th>
<th>land use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Forest</td>
</tr>
<tr>
<td>2</td>
<td>Urban</td>
</tr>
<tr>
<td>3</td>
<td>Farm</td>
</tr>
</tbody>
</table>

**Output layer**

<table>
<thead>
<tr>
<th>Geographic data</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
</tr>
<tr>
<td>A2</td>
</tr>
<tr>
<td>A2</td>
</tr>
</tbody>
</table>

**Attribute data**

<table>
<thead>
<tr>
<th>ID</th>
<th>land use</th>
<th>soil_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Forest</td>
<td>Evard loam</td>
</tr>
<tr>
<td>A2</td>
<td>Urban</td>
<td>Evard loam</td>
</tr>
<tr>
<td>B1</td>
<td>Forest</td>
<td>Cecil clay</td>
</tr>
<tr>
<td>B3</td>
<td>Farm</td>
<td>Cecil clay</td>
</tr>
</tbody>
</table>
Feature numbers increase

Layer 1
- Geographic data
  - 1
  - 2

Attribute data
- ID: 1, Class: 0
- ID: 2, Class: 100

Layer 2
- Geographic data
  - 1
  - 2

Attribute data
- ID: 1, Cost: 10
- ID: 2, Cost: 5

Output layer
- Geographic data
  - 1
  - 2
  - 3
  - 4

Attribute data
- ID: 1, Class: 0, Cost: 10
- ID: 2, Class: 100, Cost: 10
- ID: 3, Class: 0, Cost: 5
- ID: 4, Class: 100, Cost: 5
Vector Overlay

• Continue vectors next week!
Distribute Role Assignments

- Read Merry et al. 2007 – Thursday discussion