Lecture 6: GIS Spatial Analysis

GE 118: INTRODUCTION TO GIS
Engr. Meriam M. Santillan
Caraga State University
Spatial Data

- It can be most simply defined as information that describes the distribution of things upon the surface of the earth.
- In effect, it involves any information concerning the location, shape of, and relationships among, geographic features.
Spatial Analysis

- Manipulation of spatial data into various forms to be able to extract additional and meaningful information to understand the real-world.

- Can be highly technical/mathematical but can also be very simple and intuitive.
Common Examples of Spatial Data Analysis

- Measurements (lengths, perimeters, areas)
- Querying
- Reclassification
- Buffering and neighborhood functions
- Map overlay (vector overlay and raster overlay)
- Spatial interpolation
- Trend surface analysis
- Network analysis
- Suitability analysis
Measurement in GIS

- Only an approximation of the true distance on the ground
  - Vector data are stored as line segments → even curved lines are stored as short line segments
  - Raster data are approximated using a grid cell representation
- Measurements can be stored as attributes in vector GIS
Distance Measurement

[Map showing distance measurement between two schools with a marked segment length of 19.081.21]
Distance Measurement of Raster Data

- Euclidean distance – shortest distance/path
- Manhattan distance – distance along raster cell sides from one point to another
- Proximity – concentric equidistant zones are established around the start point A, and the distance to B is taken to be the corresponding radius of the circle that intersects it
Perimeter Measurement of Raster Data

- Number of cell sides that make up the boundary of the feature is multiplied by the grid resolution

Perimeter = (number of cell sides) \times (resolution)
Area Measurement of Raster Data

- Number of cells of a feature is multiplied by the known area of an individual grid cell

\[
\text{Area} = (\text{number of cells}) \times (\text{area of individual cell})
\]
Measurements of Vector Data

- Euclidean distance – obtained by using Pythagorean Theorem
- Perimeter - sum of the straight line lengths of the boundary
- Area - sum of areas of simple geometric shapes formed by subdividing the feature of interest
Querying

- Method of data retrieval
- Can be performed either on data that are part of the GIS database or on new data produced as a result of data analysis
Two types of query for GIS

1. Aspatial Query
   - Questions about the attributes of features
   - Example: “How many elementary schools are there?”

2. Spatial Query
   - Requires information about the location of a feature
   - Example: “Where in Surigao are gas stations located?”
Combined Querying

- Individual queries can be combined to identify features that satisfy two or more spatial and/or aspatial criteria
  - Example: “Where are the municipalities which have a population greater than 5000?”

- Usually uses Boolean operators (AND, OR, NOT)
  - Example: “What are the bridges located in Surigao City? How many are they?”
Querying
Reclassification

- An important variation of querying in GIS
- Can be used in place of querying for raster GIS
- Uses reclassification rules to reclassify or assign new values to cells
- Results to a new image
- May be Boolean or Weighted
Boolean Reclassification

- Produces a two-coded image from a complex image
- Original image is reclassified to an image with only 0 or 1 as cell values
- Example, if the production forests are of interest, all cells representing forestry can be assigned a value of 1, and all other cells a value of 0.
Reclassification by Weighting

- A different weight is assigned to different feature types or classes, based on the purpose of the reclassification
- Higher weights can be assigned to priority classes while lower weights to those of lower priority
- Example, if the purpose is for forest conservation, a weight of 5 can be assigned to forestry and lower weights to other features
## Reclassification Example

<table>
<thead>
<tr>
<th>Land use</th>
<th>Original value</th>
<th>New value: Boolean</th>
<th>New value: Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forestry</td>
<td>10</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Water</td>
<td>11</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Settlement</td>
<td>12</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Agricultural Land</td>
<td>13</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Proximity Analysis (Buffering)

- Creation of a zone of interest around an entity
  - Point entity: circular buffer zone
  - Line entity: elongated buffer zone
  - Polygon entity: buffer zone has the same shape as original polygon, but larger
Sample Questions that Proximity Analysis Can Address

- How many houses lie within 100 m of this water main?
- What is the total number of patients within 10 km of this healthcare facility?
- What proportion of identified cases lies within 500 m of a suspected well (as source of infection)?
Neighborhood Functions (Filtering)

- Values of individual cells are altered based on the adjacent/neighboring cells
- May be used to remove noise/noisy pixels
Filter

- Group of cells around a target cell
- Usually called kernel or window
- Size and shape are defined by the operator
- Common shapes are square or circle
- Passed across the raster dataset (image) and used to recalculate new values for the central cell/pixel for each filter position
Topological Overlay

- Topology – relationships between linked elements
- Perhaps the most important or key GIS analysis function
- Integrating data from two or more sources to produce a new map layer
- More powerful when combined with other spatial analysis methods
Raster Overlay

- Uses the idea of map algebra ('Mapematics')
- Mathematical operators are performed on corresponding cells from two or more layers to produce an output value
- Input layers may be added, subtracted, multiplied, or divided to produce output data
- Quick, straightforward, and efficient
Raster Overlay

- An input layer is usually reclassified by Boolean to make it more meaningful and easier to analyze.
- It is important that the features on each layer are appropriately coded for easier analysis.
- Lacks spatial accuracy for coarse cells/low resolution data.
Raster Overlay
Vector Overlay

- Individual data layers to be overlaid have to be topologically correct (lines should meet at nodes and polygon boundaries are closed)
- Seldom used in isolation
  - commonly used with querying
- Time-consuming, complex, and computationally expensive
Vector Overlay

- Topology must be created for the resulting layer by calculating the intersections of lines and polygons from the input layers
  - time-consuming, especially for complex data
- Problem is the possible generation of sliver or ‘weird’ polygons from the overlay
  - results from overlaying two layers that contain the same spatial entities that are not exactly registered to one another
Types of Vector Overlay

1. Point-in-polygon
2. Line-in-polygon
3. Polygon-on-polygon
Point-in-Polygon Overlay

- used to find the polygon in which a point falls within, or find a point or points that fall within a certain polygon
Line-in-Polygon Overlay

- used to find the polygon in which a line or lines fall within, or find a line or lines that fall within a certain polygon
Polygon-on-Polygon Overlay

- used to determine which polygons from two layers intersect or are within another polygon
Types of Polygon-on-Polygon Overlay

- Union
- Clip/Cookie-cutting
- Intersect
**Union**

- Corresponds to the Boolean operator OR
- Output layer contains polygons from both input layers

This operation combines features of an input theme with the polygons from an overlay theme to produce an output theme that contains the attributes and full extent of both themes.
Clip/Cookie-cutting

- Corresponds to the Boolean operator NOT
- A polygon layer is used to cut out the portion of another polygon layer that falls within the first polygon

This operation uses a clip theme like a cookie cutter on your input theme. The input theme's attributes are not altered.

Input Theme + Clip Theme = Result Theme
Intersect

- Corresponds to the Boolean operator AND
- Output is the polygon of intersection of two polygon layers

This operation cuts an input theme with the features from an overlay theme to produce an output theme with features that have attribute data from both themes.
Spatial Interpolation

- Procedure of estimating values at unsampled sites within an area with existing observations
- Used to fill the gaps between observed data points
- Interpolated data is only an approximation of the true value
- Commonly used to construct height contours
Spatial Interpolation Methods

1. Thiessen Polygons or Voronoi Polygons
2. Triangulated Irregular Network (TIN)
3. Spatial Moving Average
Thiessen Polygons

- Used to establish area territories for a set of points
- Examples: transformation of point climate stations to watersheds; construction of areas of influence around population centers
Steps in creating Thiessen Polygons

- Draw lines joining nearest neighboring points
- Draw perpendicular bisectors of these lines
- Use the bisectors as edges of the polygons that make up the surface

Figure 1 - Thiessen polygons representing areas in which all locations are closest to each point.
Triangulated Irregular Network

- An elegant way of constructing a surface form a set of irregularly spaced data points
- Most often used to generate DEMs
- Exact interpolation based on local data points
- Adjacent data points are connected by lines to form a network of irregular triangles
Triangulated Irregular Network
Spatial Moving Averages

- Most common interpolation method used in GIS
- Involves calculating a value for a location based on the range of values attached to neighboring points that fall within a user-defined range
  - uses a filter that moves across the region of interest and calculates interpolated values per location
An approximate interpolation → does not honor the values of known data points (recalculates them)

Used when values of known data points are not exact/erroneous

Examples: interpolation of census data, questionnaire responses, soil PH, rainfall infiltration rates
Surface Analysis

- Slope and aspect calculation
- Visibility analysis
Visibility Analysis

- Identification of areas of terrain that can be seen from a particular point on the terrain surface

- Applications:
  - Locating towers for geodetic surveying
  - Locating radio transmitters for maximum coverage
Network analysis

- Involves linear features or network of linear features that are topologically structured → roads, rivers, pipelines, cables, etc.

- Examples:
  - shortest path
  - time of travel
  - route tracing
## Network Analysis Example

<table>
<thead>
<tr>
<th>Directions</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Start out going East on HENLEY ST towards WARREN ST.</td>
<td>0.1 miles (0.1 km)</td>
</tr>
<tr>
<td>2: Turn RIGHT onto WARREN ST.</td>
<td>0.0 miles (0.1 km)</td>
</tr>
<tr>
<td>3: Turn RIGHT onto CHELSEA ST.</td>
<td>0.0 miles (0.1 km)</td>
</tr>
<tr>
<td>4: CHELSEA ST becomes CHELSEA ST/CITY SQ.</td>
<td>0.1 miles (0.1 km)</td>
</tr>
<tr>
<td>5: Turn RIGHT onto CITY SQ/NEW RUTHERFORD AVE/SR-99 N.</td>
<td>0.0 miles (0.1 km)</td>
</tr>
<tr>
<td>6: Stay straight to go onto NEW RUTHERFORD AVE/SR-99 N.</td>
<td>0.2 miles (0.3 km)</td>
</tr>
<tr>
<td>7: Turn SLIGHT LEFT onto SR-99 N</td>
<td>0.4 miles (0.6 km)</td>
</tr>
<tr>
<td>8: Turn SLIGHT LEFT onto SR-99 N/RUTHERFORD AVE.</td>
<td>0.1 miles (0.1 km)</td>
</tr>
<tr>
<td>9: Turn SLIGHT LEFT onto SR-99 N</td>
<td>0.3 miles (0.4 km)</td>
</tr>
<tr>
<td>10: Turn SLIGHT LEFT onto SULLIVAN SQUARE OPAS</td>
<td>0.4 miles (0.7 km)</td>
</tr>
<tr>
<td>11: Turn SLIGHT LEFT onto MYSTIC AVE</td>
<td>0.7 miles (1.1 km)</td>
</tr>
<tr>
<td>12: MYSTIC AVE becomes MYSTIC AVE/SR-38 N</td>
<td>1.2 miles (2.0 km)</td>
</tr>
<tr>
<td>13: Turn LEFT onto HARVARD ST</td>
<td>0.6 miles (1.0 km)</td>
</tr>
<tr>
<td>14: HARVARD ST becomes WARNER ST.</td>
<td>0.2 miles (0.3 km)</td>
</tr>
<tr>
<td>15: Turn RIGHT onto POWDER HOUSE SQ.</td>
<td>0.1 miles (0.1 km)</td>
</tr>
<tr>
<td>16: Turn RIGHT onto BROADWAY</td>
<td>1.0 miles (1.6 km)</td>
</tr>
<tr>
<td>17: Turn LEFT onto ALEWIFE BROOK PKWY/SR-16</td>
<td>0.4 miles (0.7 km)</td>
</tr>
<tr>
<td>18: ALEWIFE BROOK PKWY/SR-16 becomes ALEWIFE BROOK PKWY/SR-16/US-3</td>
<td>0.4 miles (0.7 km)</td>
</tr>
<tr>
<td>19: Take CONCORD TURNPIKE/SR-2 W.</td>
<td>4.7 miles (7.6 km)</td>
</tr>
<tr>
<td>20: Take the WALTHAM ST. exit, exit number 54B, towards LEXINGTON</td>
<td>0.2 miles (0.3 km)</td>
</tr>
<tr>
<td>21: Merge onto WALTHAM ST.</td>
<td>1.9 miles (3.0 km)</td>
</tr>
<tr>
<td>22: Turn RIGHT onto MASSACHUSETTS AVE/MASS AVE/SR-225</td>
<td>0.0 miles (0.0 km)</td>
</tr>
</tbody>
</table>

**Total Distance:** 12.9 miles (20.8 km)  
**Estimated Time:** 24 minutes
Thank you!