Spatial Analysis: an overview

Introduction

The next two lectures mark an important shift in this class. We have covered database development and data acquisition, and now turn our attention to data analysis and display. The next two lectures will cover basic spatial analysis techniques. We will begin by looking at definitions of models, with particular emphasis on spatial analysis as a type of modeling; following this, we will begin an overview of different types of spatial analysis. The concentration will be on non-statistical, vector modeling, although there will be some discussion of raster modeling where appropriate.

Model Definitions

Spatial analysis is an attempt to model both physical objects and processes. These processes can be either natural or human creations that take place in the world around us. To the right are four different definitions of Model. The common factors for all four definitions are that models are abstractions of the real world and real world processes. In a GIS, these abstractions are digital representations of real world objects and information about the objects (georelational), and the processes are represented by a variety of operations on the objects. This type of process modeling comes in a number of forms that fall into two basic categories: cartographic modeling and map algebra.

- Attempt to duplicate nature in order to simulate or provide new information on processes (Wheeler 1988)
- An artificial construction in which parts of a source domain are represented by means of a structure preserving function in a target domain (Worboys 1995)
- An abstraction and description of the real world or part of it (McDonnel and Kemp 1995)
- A set of algorithms written in computer code that describe a given physical process or natural phenomenon of the earth’s surface (Burrough and McDonnel 1998)

Cartographic modeling creates a set of interacting ordered map operations that act on raw data as well as derived data to simulate a spatial decision-making process. Cartographic modeling can be carried out in either or both raster and vector data structures. Good examples of cartographic models are suitability/capability models, like the one you have been creating for Lab 11. These models involve processes operating on data in an ordered and interactive manner. The order of the operations is particularly important, and it is not possible to go from stop A to step E without passing through steps B, C, and D.
Map Algebra is a different type of modeling, involving the use of algebraic equations to operate on raster data in the same way that they work on single numbers. For example, the equation GRID A + GRID B = GRID C. In this instance grids A and B are summed on a cell by cell basis to create grid C. Although in the equation the grids are written as if they are single numbers, they represent as many numbers as there are cells in the raster, often millions of individual numbers. Map algebra can range from very simple to very complex equations. Although they are typically thought of as local operations, they can be expanded to neighborhood or zonal operations. Map algebra is not covered in the labs for this course, but in the advanced GIS courses it is covered in depth.

An Overview of Spatial Analysis Categories

The scope of spatial analysis operations is larger than can be covered in two lectures. Because of this, we will look primarily at non-statistical vector analysis, although raster will creep in occasionally. We will look at six broad categories of modeling (below), with examples of each. The categories are arranged from simplest to most complex. The first two will be covered in the first lecture, and the remaining four in the second lecture. This discussion has been informed by Andy Mitchell’s book\(^1\) on GIS analysis. If you are

---

interested in examining the topics in greater detail, it provides an excellent starting point. The six categories we will be discussing are:

- Mapping where things are
  Finding simple patterns in your data
- Mapping most and least
  Using quantities to find patterns in your data
- Mapping Density
  Using concentrations to find patterns in your data
- Finding what’s inside
  Discovering whether an activity occurs inside an area
- Finding what’s nearby
  Discovering whether an activity occurs near a feature
- Mapping change
  Using change over time to find patterns in your data

**Mapping Where Things Are**

Objects are often mapped in order to find patterns in data; hopefully patterns that will help researchers understand the phenomena being mapped. This simple form of analysis requires both geometry – points, lines or polygons – and attributes. The GIS is necessary for storing and displaying geometry and attributes. The user is responsible for sub-setting data, determining classes, and assigning symbols for data display, and for identifying patterns using ocular review of the map.

In the map to the right, migrant deaths in southern Arizona are symbolized as red dots. Simple maps like this can be powerful story tellers that allow the user to identify patterns in data. For example, note the very dense cluster of deaths to the southwest of Tucson, and the string of deaths along I-10 between Tucson and Phoenix.
Sub-setting and classification of data are also important components of this type of spatial analysis. In the map to the right, a subset of migrant deaths in the area southwest of Tucson has been created. It has also been classified by cause of death. Notice that exposure is the most common cause of death, and that death by accident usually occurs along roads.

**Mapping Most and Least**

Mapping features based on quantity provides another layer of information beyond location. It can add meaning to patterns and even correct misunderstandings about patterns based strictly on the location of objects. Mapping most and least requires both geometry and quantifiable attributes. These are stored in the GIS, and manipulated by the user to create maps of attribute quantities. As with mapping location, the user’s primary task is to make decisions about the data that will tell the story in the best way possible. When mapping most and least, this is primarily about data symbolization. Because it is quantifiable, the user must make decisions about data normalization and classification, and about the type of symbolization that best fits the data and the story that will be told.

Colors and color ramps are important devices for quantitative mapping. In general, color ramps that use cool colors for low values and hot colors for higher values are most effective. In the map to the left, population by census block are ramped between green and red. Green blocks have smaller populations than red blocks. In monochromatic ramps, lighter colors usually equate to low values and darker colors to high values.
Normalization of data is also an important consideration when mapping quantities. In the maps above, Spanish speaking households in Tucson are mapped. On the left, total per block is shown, while on the right these data have been normalized as percent of total households. The stories that these maps tell are different in important ways.

For point data, mapping quantities can be done in a number of different ways. The maps above are related to an archaeological survey that collected surface pottery from a random sample of locations within a circular project area. The map on the left shows the plots where pottery was collected; on the right the same plots are shown as graduated symbols based on the amount of pottery collected from each plot. Clearly there was more pottery, and presumably more people in the southeast part of the project area.
Another way that quantities associated with points can be mapped is by using contours. In the map to the left, the pottery collection points appear as blue marks. The quantities of potter collected at each plot are represented by contours interpolated from the points.

When working with point data, the use of rasters to show quantities is also effective. In the map on the left, quantities of pottery collected from the sample points has been interpolated. A monochrome color ramp has been employed with light to dark colors matching values ramped from low to high.

Mapping Density

Density mapping helps the user to discover concentration of features. It is similar to mapping least and most, but normalizes these data by area. In order to map density it is necessary to have geometric features, an area measure, and quantities (either counts of features or values in an attribute. The GIS stores and displays quantities per area measure. The user plays a big role in density mapping, starting with determining the type of density method that will be used; either mapping by an already defined area or by developing a density surface. Users must also determine whether they will map feature counts or feature values as well as determining symbolization.
Determining density for shaded polygons requires both area and a quantitative attributes. Density is measured by normalizing the quantitative attribute by an area attribute for each polygon. In the example to the right, asthma cases per square kilometer are mapped.

Here the same data is displayed as a dot density map. For dot density maps, area is implicit, it is the size of the polygon, and density is measured as dots per polygon. Higher density will cause polygons to fill with dots, whether they are small or large. In this particular map, each dot equals 10 diagnosed asthma cases.

Calculating density for points or lines requires the user to apply an area measure, either by using a raster or by creating a polygon layer.
In this map, area is delineated by a raster with a cell size of 20 X 20 meters. The red marks are archaeological sites, and the density surface represents a count of sites within 1000 meters of each raster cell.

Here the user has created an array of one square kilometer polygons. Density was determined by calculating the length of all road segments in each polygon.

Finding What’s Inside

Finding what’s inside allows the user to monitor what is occurring in an area, or to compare multiple areas based on what’s inside of each. Geography for multiple feature classes, at least one of which must be polygonal, along with attributes are necessary to determine what is inside of areas. The GIS stores and displays geography and attributes and performs overlay operations. The main decision that the user must make regards the type of operations to perform. These operations range from simply drawing areas and features for visual evaluation, to selection by location, to analytic overlay operations.
Simply drawing areas and features does a good job of discovering if features are inside or outside of an area. This works with points, lines, polygons, or surfaces. Because it is visual only, the user can’t do anything other than look at the map.

Selection by location allows the user to leverage the spatial abilities of the GIS to select features. It is an excellent method for making a list, or summary table of features inside an area. It works with points, lines, or polygons, but it can’t tell you which objects are in each particular area.

Selecting by location comes in a variety of forms, including:

- Intersect
- Contain
- Are contained by
- Share a line segment
- Touch boundary
- Are identical
- Are centered in
- etc…

In this map, vegetation polygons were selected if they intersected the polygon representing Fort Huachuca (the red line). Selected features can then be listed, or summarized (below)
Feature overlay is the most sophisticated, and processor intensive of the different methods used for finding out what’s inside. It is the only method discussed here that can be used to determine which features are in which area. It works with points, lines, areas, and surfaces. The most common overlay operations include union, intersect, identity, and difference.

Identity is one of the most useful of the overlay operations. It allows the user to pass the features in one layer through polygon features in another layer. As they pass through they pick up the attributes of the other feature class. In the maps below, the identify operation passes a point layer (fire_pts), through a polygon layer, (admin). As seen in the tables below, the new point layer contains the attributes of both original layers.
Finding What’s Near

This form of analysis allows the user to discover what’s occurring within a set distance of a feature. It is often used for determining travel distance or cost. Source and target layers are needed for this analysis to work. The GIS stores the geometry and attributes, and calculates distance from features. The user’s job is to determine which operation should be used to calculate distance: straight-line distance, distance/cost over a network, or distance/cost over a surface.

Calculating straight line distance in a vector setting requires the user to create buffers around features. These buffers can then be used to discover which features are within a particular distance of the buffered feature/s. In the image to the right, coyote dens have been buffered by 3000 meters.
Determining distance/cost over a network works with both points and lines. It gives precise travel cost over a network, but requires a very accurate network layer. In this map, network distance is used to determine sales zones for retail outlets.

A distance surface uses a continuous raster to measure overland travel distance or cost. It measures Euclidean distance, cost distance, or path distance, allowing the user to combine several data sets to measure absolute distance. In this map, path distance from an archaeological site is being measured. Path distance is a technique that measures true distance (up hill and down hill) and allows the incorporation of friction surfaces to calculate cost distance across a surface.

**Mapping Change**

GIS allows the user to map how things change over time, or how things move. Knowing how things change allows greater understanding of how objects behave and researchers can begin anticipating future events. Geometry and/or attributes that change over time are necessary to map change. The GIS facilitates the display of data that change over time in location, character, or magnitude. The user determines which type of change is most appropriate for the data (time series, tracking map, or measured change), and creates display parameters for the GIS to follow.

**Time Series**

Time series are multiple maps that measure movement or change in character of an object of phenomena. Time series include trend, cycle, and before and after maps. These are characterized by strong visual impact, but because they don’t indicate exact moments of change users need to make careful visual comparison of maps. In the example below,
development on Cape Cod is shown in a series of three maps. Beginning in 1951 the maps stop in 1985 and predict build-out of the cape by 2025.

Tracking Maps

Tracking maps measure movement of objects across the surface of a map. When concentrating on a few objects, they are easily interpreted maps for distance and rate of travel; however, as the number of objects being tracked increases, interpretation of these maps becomes more difficult. In this example, Adelita the loggerhead turtle was tracked crossing the Pacific Ocean from Baha to Japan over the course of several months.

Measured Change

Measuring change in maps requires quantitative data collected at two different time periods. Change is measured by subtracting one from the other. These maps are good because they show actual differences in value. In the maps below, the change in nighttime to daytime population in San Diego is shown.
Summary

The past two lectures have examined spatial analysis. We covered model definitions, and how these definitions relate to spatial analysis. The bulk of our time was spent on an overview of different types of analysis. We looked at:

- Mapping where things are
- Mapping most and least
- Density mapping
- Finding out what’s inside
- Finding out what’s near
- Mapping change

Please note that this was a simple overview, and that there is much more that could be said about spatial analysis. If you are interested in learning more about this topic, there are many good books and articles, as well additional GIS courses you can take.