Spatial Modeling – output format and Generation

Praveen Kumar, *Department of Geography*

Basic of Spatial Modeling

Spatial Modeling in GIS allows us to turn data into information and create new data (derivative datasets) by manipulating existing spatial features and their related attributes. GIS packages are equipped with a variety of analysis functions that allow us to manipulate both vector and raster data formats. These functions can be thought of as a set of tools for spatial analysis, and in fact several GIS applications use this "toolbox" analogy in describing the geoprocessing functions available. Tasks performed by a GIS analyst will typically involve making use of several of these analytical tools. For example, a simple analytical problem might be to determine the amount of agricultural activity within 500m of streams, perhaps as a means of quantifying riparian disturbance, or water quality degradation. To answer such a question, an analyst must buffer the streams by 500m, overlay the buffers with agriculture land use polygons, and then quantify the resulting intersection. This module seeks to enumerate and define many of the common analysis tools which are available in most commercial GIS applications, which can then be combined to resolve specific analytical problems.

For organizational reasons, these functions are divided into the following topics:

1. Vector analysis methods
2. Raster analysis methods
3. Generalizing data

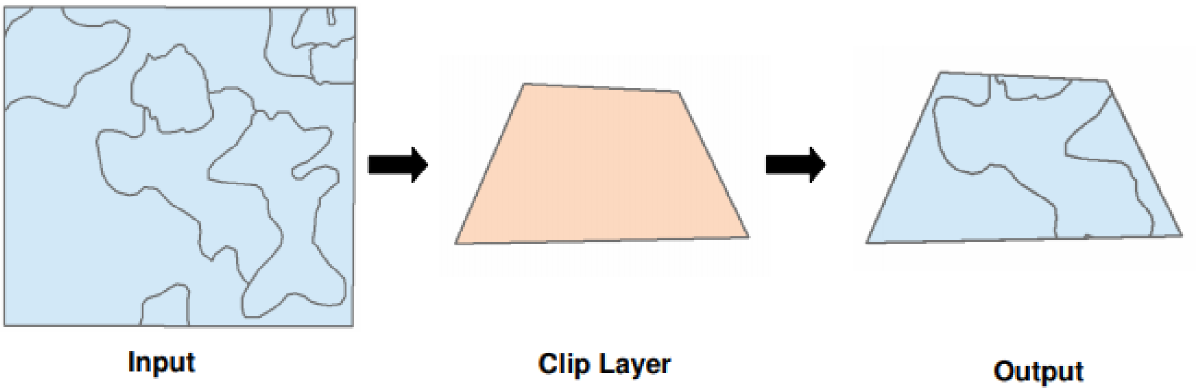
Vector Analysis Method:

Extraction

Extracting portions of data is an effective means of isolating specific areas for further processing or data analysis. Similar to queries and selection sets, extraction functions can be used to reduce the size of datasets and/or facilitate more complex interpretation. While the development of queries and selection sets also will allow you to isolate portions of a dataset, extraction techniques differ in that these portions of data are isolated in a permanent way - through the creation of new data layers. GIS software packages provide a suite of tools to extract data, the most useful being, clip, select, split and erase.

Clip

Working much like a cookie-cutter, clip allows you to intersect two feature layers to extract a portion of a dataset (the input layer) based on the spatial extent of another dataset (the clip layer). The clip function creates a new data layer (output) consisting of the features of the input layer that fall within the extent of the clip layer Clip is useful for developing a subset of features from a series of existing data layers to match a common boundary, for example the boundary of a study area or a jurisdictional boundary (e.g., a province, county, state, or municipal boundary). For example, an urban planner might wish to look at a street network layer, but only those streets falling within a certain municipal boundary.

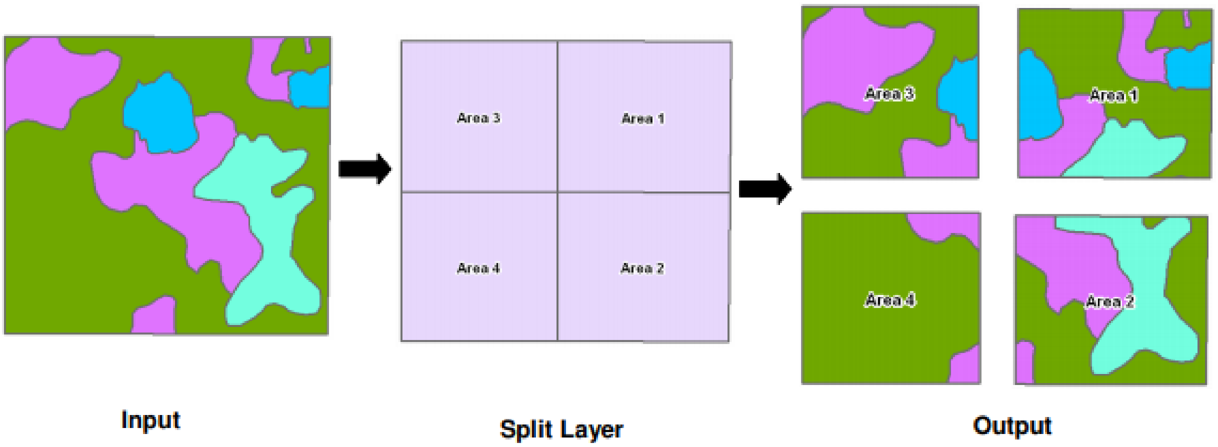


Clip Example

Clipping would be useful in order to permanently extract the street features matching the extent of the municipal boundary. The input layer to be clipped may contain points, lines or polygons; however, because the element of area is required, the clip layer must be a polygon. The field names and attributes of the features in the output layer's table are maintained (i.e., they are identical to those of the input table). One potential exception to this rule are area, length and perimeter fields, which, depending on the software and/or data format being used, may or may not automatically recalculate. The values of any features intersected by the clip boundary may require updating to reflect the change in area.

Split

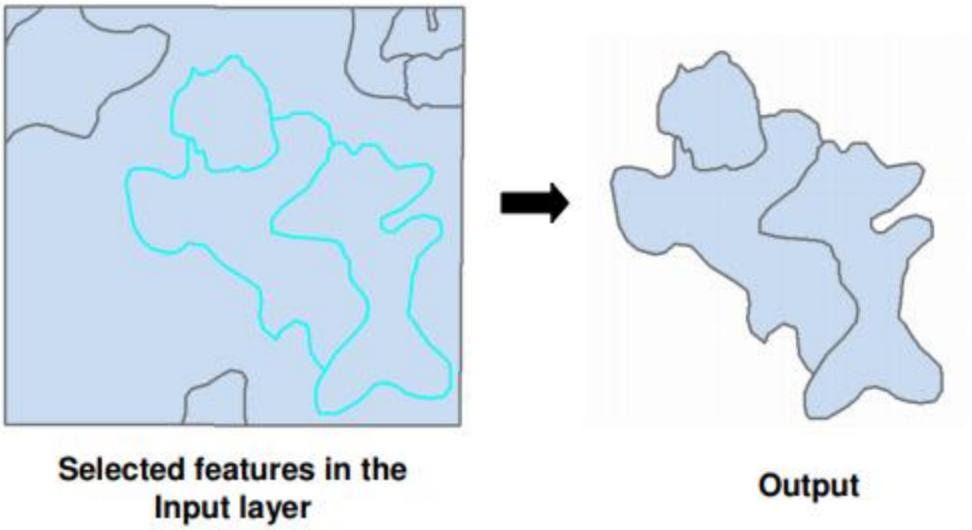
Split is used to divide an input layer into two or more independent layers, based on geographically corresponding features in a split layer. Similar to the clip function, the input layer may consist of point, line or polygon features, however, the split layer must be a polygon to define the areal extent of the analysis. The features in the input layer are broken up along the boundaries of the split layer features as illustrated in Splitting is essentially a means of simultaneously executing a series of clips along boundaries defined by the split layer. The number of output layers is dictated by the number of split features (i.e., if the split file contains four polygons, the input file will be split into four separate files). Each resulting output layer will be named with the unique attribute value present within the selected field from the split layer. As with clipping, the field names and attributes of the input table are maintained in the output layers. The split function would be useful for dividing a large coverage into jurisdictional areas, for example, the zoning data associated with a city could be divided based on municipal boundaries or a national map series could be developed by dividing topographic data based on a defined grid.



Slit Example

Select

The Select tool may be used to create a new layer containing features extracted from an input layer. This is achieved through the execution of a user-defined query expression to select a subset of the data; these selected features are permanently extracted to a new output layer



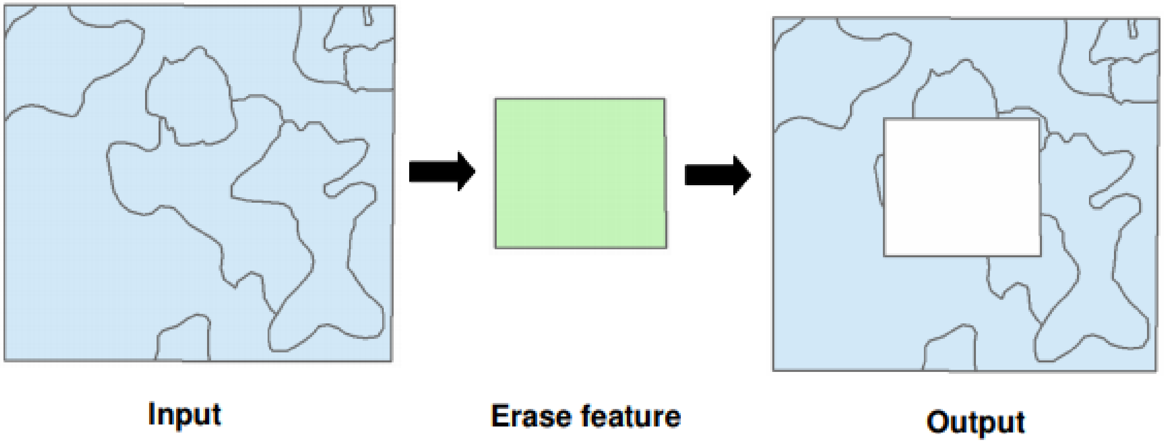
Select Example

To build on the example above, the urban planner might wish to look at only double-line streets in the particular municipality of interest. In this case, he or she would execute a selection query to extract only those desired features to a new layer.

Erase

Erase creates a new output layer by discarding features from the input layer that fall within the area extent of the erase (overlay) layer (Figure 4). The input layer can be points, lines, or polygons; however, because the element of area is required, the erase layer must be a polygon. Features in the output layer will be of the same geometry type as the features in the input layer. Examples of how the erase function could be used include:

In a map layout, erase can be used to develop a mask to allow only those features falling within a given area (e.g., a study area boundary) to be displayed. In a suitability analysis, erase could be used to apply suitability rules. For example, if potential sites have to have a 200 metre setback from wetlands then wetland features can be buffered by 200 metres and the buffer polygon used as the erase layer to remove potential sites falling within this zone from consideration.



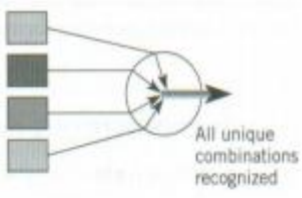
Erase Example

Overlay

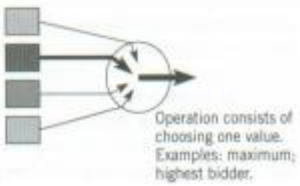
Central to GIS analysis is the integration of data to reveal the relationship(s) between two or more data sources. Overlay is one method of integrating information as it combines the spatial and attribute data from two or more input layers to create a new output layer. The spatial form of the new layer is shaped by the geometric intersection of the input and overlay features. Generally, the overlay of two or more layers results in a more complex output layer, with more polygons, intersections and/or line segments than what is present in the input layers. Each feature in the newly created output layer contains a combination of attributes from the input layers. Overlay functions, when associated with geometrical (or 'physical') overlays of data layers, are implemented by certain mathematical operations – both arithmetic and logical. Arithmetical operations commonly used, but not limited to, include addition, subtraction, division and multiplication. Logical operations are aimed at finding where specified conditions occur and use logical operands such as: AND; OR; >; and <

As discussed later in this module, methods for overlaying vector data differ from those of raster data related methods. However, basically vector-based methods do not generalize the data but, due to relatively more intensive processing requirements, may be more appropriate for smaller or sparser datasets. Raster analysis methods generalize the data based on the largest cell size found among the input layers. Raster-based grid calculations are, however, often faster and easier. Four basic rules for combining the attributes of several layers can be applied to overlay analyses, and these are presented in Figure 5Figure below. While these rules are presented here with the vector analysis methods, they are equally relevant to a discussion of raster overlay methods.

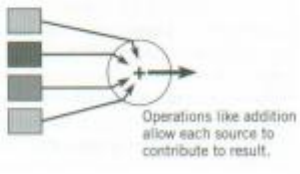
1. Enumeration Rule: Each attribute is preserved in the output layer and all unique combinations are recognized. For example, a soils layer, vegetation layer and precipitation layer are overlaid yielding a derivative coverage with a unique polygon for each possible combination.



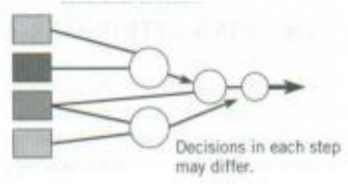
2. Dominance Rule: One value wins – the dominant (e.g., highest) value is the only one value assigned. For example, an overlay based on a series of sensitivity layers would assign the highest sensitivity value to each derivative polygon.



3. Contributory Rule: Each attribute value contributes to the result - each source layer contributes to the result. For example, environmental sensitivity could be calculated based on the sum of a set of input layers: wildlife habitat sensitivity; riparian sensitivity; slope; and proximity to human disturbance.



4. Interaction Rule: A pair of values contribute to the result (i.e., decisions in each step may differ)

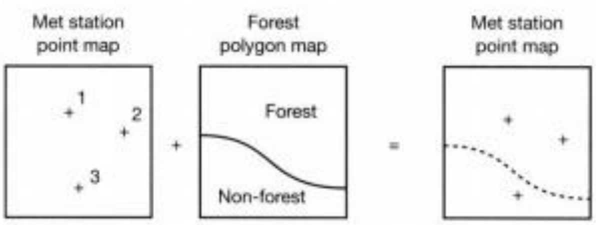


Overlay Rules:

Three main types of vector overlay exist:

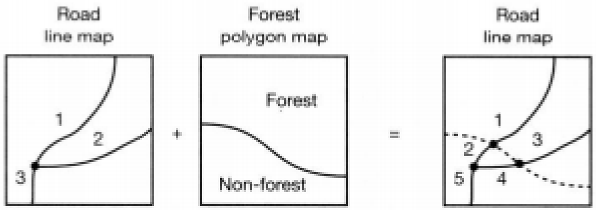
point-in-polygon – The point features, which maintain their spatial location and attribute integrity in the output layer, are also assigned the attributes of the polygon they fall within.

This type of overlay might allow an association between meteorological stations (the met station point layer) and vegetation types (the forest polygon layer) to be identified (Figure 6). The output layer would be the meteorological station point file with the addition of a vegetation type attribute.



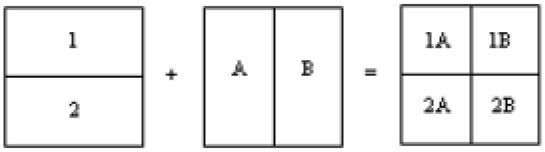
Point in Polygon overlay example

line-in-polygon – The line features, which maintain their spatial location and attribute integrity in the resulting output layer, are assigned the attributes of the polygon they fall within. This type of overlay would allow you to determine the vegetation types (derived from the forest polygon layer) associated with each line of a road layer (the road line layer). Because a line may overlap multiple polygon types, the output layer (the road line map) will generally contain more line segments than the input layer.



Line in Polygon overlay example

polygon-on-polygon – The polygon geometries from the input and overlay layers combine to create a new set of polygons where each new polygon maintains the attributes from both input layers (Figure 8). This type of overlay might be used to find the association between slope and avalanche chutes. Polygon-on-polygon is the most common of the vector overlay methods.



Polygon on Polygon overlay Example