GEO/EVS 423 / EVS 523
Exercise 2: Geoprocessing

The basic operations of GIS are collectively known as geoprocessing. In this exercise, you will examine and manipulate maps of various features of northeast Ohio. Copy the NorthEastOhio.gdb geodatabase from the P: drive to your X: drive and open it in ArcGIS. Add all of the layers except for the neowsheds arcs, tics, and points. These layers are various water-related and land-cover layers for northeastern Ohio, as well as an image of county boundaries. The counties shown in the image are Lorain, Medina, Cuyahoga, Summit, Lake, Geauga, Portage, Ashtabula, and Trumbull. The layers are hydrologic points, lines, and polygons, land-cover areas, watersheds, and county boundaries. You should take some time to examine each layer and determine what it includes. The geodatabase also includes most of the metadata files for the various layers; these are web files that can be opened in Firefox.

Your first step is to examine the metadata for the various layers. When you open the attribute tables for some of the layers, you will notice that some of the layers have labels and others do not. You should add a class name column to any attribute table that does not have one and add the appropriate class names (You do know how to do that, don't you?).

Next, you should take a good look at each layer and create a map that shows the land cover and the hydrology of northeast Ohio most effectively. You should try to make a map that conveys the maximum amount of information, complete with legends and text. Print this map.

Now, let us concentrate on Cuyahoga County. Turn on the labels (if you don't already have them turned on) on the county map so that the counties are named. Now select Cuyahoga using the Select by Attributes dialog box. To do this, open the attribute table for the counties layer. Look at the names of the fields in the table. What field includes the county name? Click on Options -> Select by Attributes. Click on the name of the field containing the county name in the fields box and then the = button. Click on the Unique Values button (this is important). The names of the counties appears. Double-click on Cuyahoga to choose it, so that the SQL statement finally reads, "SELECT * from NEOhio WHERE "Name" = 'Cuyahoga'. Hit Apply and then Close.

It is useful to understand the syntax of this statement: Select Everything (the * = everything) from the NEOhio table where the value in the Name field (that's why Name is in double quotes) is equal (i.e. the =) to the string Cuyahoga (that's why Cuyahoga is in single quotes). This is the standard way in which an SQL, or Structured Query Language, statement is made, and you will be using SQL a lot in Arc.

You can save Cuyahoga as a feature class all by itself by right-clicking on the NEOhio listing in the Table of Contents window, then clicking Selection -> Zoom to Selected Features. Then right-click again on the NEOhio layer, then click Data -> Export Data. The output features class should be in the NorthEastOhio.gdb geodatabase and named in such a way as to recognize that it refers only to Cuyahoga. Add the new layer to your map layout.

You might wonder if there's an easier way to do this. If you had a very complex layer with hundreds of features, the answer would be no. In this case, though, there is an easier way. Go back to the full extent of your map and click on Geauga in the NEOhio layer. Now right-click that layer and then click Data -> Export Data and save your map in the NorthEastOhio.gdb geodatabase as Geauga.

Spatial Selection and Clipping of Data

The NorthEastOhio.gdb geodatabase contains several layers dealing with aspects of hydrology and land cover in northeast Ohio. Let's look at rivers in Cuyahoga County. The feature class NEOHydroLines shows rivers all over northeast Ohio. You can limit the rivers on your map to those in or intersecting with Cuyahoga County by clicking Selection (on the Arc main menu) -> Select by Location. Choose "Select Features From" as the selection method, NEOHydroLines as the target layer, Cuyahoga (or whatever

you named your layer for Cuyahoga County) as the source layer, and "Target layer(s) features intersect the source layer feature" as the spatial selection method. Be sure that you understand what all of these fields mean. The notion of "source" and "target" deserve some special explanation, as they aren't obvious. The "source" layer is the layer that is used as the basis of the spatial selection. In this case, it's the Cuyahoga County. And it's not the representation of the county in the NEOhio layer; it's the separate Cuyahoga layer. It's important that the source layer be a polygon layer that comprises no more than or no less than the spatial bounds that you wish to consider. The "target" layer is the layer that is being examined through the spatial bounds represented by the source layer. If you will, the source layer is a cookie cutter; the target layer is the cookie dough. The result of the spatial selection is that portion of the dough that was selected by the cookie cutter. By choosing the selection method of target features intersecting . . ., we get any rivers *within* the county as well as those few rivers that extend into adjacent counties.

When you have completed the selection by location, turn off the NEOHydroLines layer so that you can verify that only the Cuyahoga rivers have been selected. Save the Cuyahoga rivers as a separate feature class within the NorthEastOhio.gdb geodatabase.

You will notice that some of the selected rivers by the "Select by Location" tool extend outside of the Cuyahoga County boundaries. Can you figure out why? How far do they extend? That might be useful for some analyses, but you might to limit your image specifically to the boundaries of Cuyahoga County. Now click on Geoprocessing (from the main Arc menu) -> Clip. Use NEOHydroLines as the input feature and Cuyahoga (or whatever you called your Cuyahoga layer) as the Clip Feature. Give your output feature a suitable name – but not the same as the Cuyahoga river feature you created above – within the NorthEastOhio.gdb geodatabase.

Compare the two maps of Cuyahoga rivers you made – first by spatial selection and the second by clipping. Are there differences? If so, what are they? Can you tell how they came about?

Merging and Dissolving Features

Dissolving is a relatively simple concept. It removes boundaries separating objects when these objects are characterized by a characteristic (i.e. a field in the attribute table) that is the same for the adjacent objects. Open the NorthEastOhio.gdb geodatabase and look at the Watersheds layer, NEOWsheds. Open the attribute table, and consider the HYUN field; this shows the Hydrologic Unit Number for each watershed. On the main Arc menu bar, click on Geoprocessing -> Dissolve. Choose NEOWsheds for the input feature, and use HYUN as the dissolve field. Give your output feature class an appropriate name. For the statistics field, choose the feature ID as the Statistics Field and choose COUNT as the statistics type. The resulting statistics field will show you how many sub-watersheds are part of each hydrologic unit. Using the Identify tool (the blue circle containing the "I") what can you tell about each of the major watersheds identified by the Dissolve?

Carry out the clipping operation you did for Cuyahoga County rivers on Geauga County. In a new map, add the resulting river maps for Cuyahoga and Geauga counties, Save this as a map. Now click Geoprocessing -> Merge. Add the two layers as input datasets, and give an appropriate name to the output dataset (in the NorthEastOhio.gdb geodatabase). Print the resulting map.

Intersection and Union of Layers

Intersecting in GIS is basically like using a cookie cutter. The NorthEastOhio.gdb geodatabase includes layers which you have created showing watersheds of the major river systems in northeast Ohio. Identify the watersheds of the Cuyahoga and Grand rivers. Select each and save the watershed as a feature so that you have two layers, one of the Cuyahoga and the other of the Grand.

From the main Arc menu bar, click Geoprocessing -> Intersect. You will intersect each of the watershed features with the NEOLandCover layer. For each, the two input features are your watershed feature and

NEOLandCover. For each, give the output feature class an appropriate name, and make sure that the output type is Polygon. The result will be a land-cover image of each watershed. Open the attribute table for each, and summarize the total area in each land-cover type. Remember that you do that by right-clicking on the column selector for the land-cover class and choosing Summarize. When the summarize window opens, choose shape_area -> sum as the summarization statistic. Add the resulting table to your map. The format of this table is *.dbf, which means that you can read it into a spreadsheet such as Excel or LibreOfficeCalc. While you are at it, you should also make a summary table for land cover for all of northeastern Ohio. When you have all 3 tables, open the spreadsheet of your choice, and read in each of the three tables. Convert the area from square feet to square miles, and use Sum() to calculate the total area of the watershed or area. Now calculate the proportion of the land in each watershed occupied by each land-cover type.

Imagine, for the moment, that you want to present a comparison of land-cover in each of the watersheds. Create a layout that includes an appropriate image of the land cover in the watershed as well as the summary statistics and explanatory text. Print this map.

When you intersected the watershed features with the land-cover map, you used the watershed as a cookie-cutter to extract information from the regional map. Now take the two watershed-level land-cover maps and put them together. Click on Geoprocessing -> Union. The two watershed-level maps are your input features; your output should have a suitable name and be put into the NorthEastOhio.gdb geodatabase. You will be uniting ALL attributes. The resulting map will be added to your map. Set the symbology to show land-cover class – but something appears wrong. What is it? Do you know how to "fix" it so that all land covers for both watersheds can be shown?

Buffering

Proximity is important – What areas are near or adjacent to other areas of interest? One of the areas of natural interest in northeast Ohio is the Cuyahoga River. You will determine the land cover in one or more buffer zones around the Cuyahoga and its tributaries in various portions of the watershed.

Use your polygon image of the Cuyahoga River watershed to clip the NEOHydroLines image, so that you have an image of all of the rivers within the Cuyahoga drainage. Also, clip the neowsheds_polygon image, so that you have an image of the subwatersheds within the Cuyahoga drainage.

Now create a buffer zone around the streams in the watershed. Use Geoprocessing -> Buffer to open the Buffer dialog box, with your clipped river image as the input file. Choose a reasonable distance (e.g. 100 feet or so) as a buffer distance, and give your output file an appropriate name. When you hit OK, ArcGIS will compute a buffer area around the permanent streams in the Cuyahoga drainage.

You can then find the land cover in this buffer zone by intersecting it with the land-cover map of the watershed you created earlier. Summarize the total areas of each land-cover type in the buffer zone for the entire watershed. Export the data table.

Now intersect the land cover in the buffer zone with your clipped subwatershed image. Open the attribute table for the resulting image, and sort it by the subwatershed name. Choose 3-4 different subwatersheds (you can figure out which ones you want to choose by looking at the subwatershed polygon file), and summarize the areas of each land-cover type within the buffer zone in the subwatershed. You do this just as you did for the watershed as a whole, except that you summarize only the selected records. Export the data tables, and select the subwatersheds they represent. Print a map showing the subwatersheds you've considered and the tables showing land-cover types comprising them. Does this sort of analysis suggest approaches to understanding the sources of pollution or other use impairments in natural resources such as rivers?

Portfolio

You should print out the following maps from this exercise. Be sure that your *name* and the *printout number* appear on each map. If you want, you can submit more than one map for a given printout number.

2-1 Your map of the land cover and hydrology for northeast Ohio, with legend showing land-cover types

2-2 Your map of streams in Cuyahoga County with two data frames, one showing the "select by location" streams, the other showing the "clip" streams. You may also want to add a text field indicating why there's a difference.

2-3 Your map of the streams in Cuyahoga and Geauga counties as you've merged them

2-4 Your maps of land cover in the Grand and Cuyahoga River Watersheds and all of northeast Ohio, with a table showing percentages of each land-cover class in each watershed

2-5 Your map of land cover in the merged Grand and Cuyahoga River Watersheds.

2-6 Your map showing subwatersheds in the Cuyahoga River and tables of land-cover in several of those subwatersheds.