University of Massachusetts Boston – Department of Environmental, Earth, and Ocean Sciences

EEOS 383 - GISCIENCE FOR WATER RESOURCES RESEARCH – Spring 2010

Instructor: David E. Tenenbaum

Exercise 01: Vector to Raster Conversion and the Raster Calculator

Introduction

While you probably are familiar with geographic information systems, chances are good you have worked primarily with spatial data that used the vector data model. In this course, we will be primarily making use of data and techniques that use the raster data model. In this first exercise, you will familiarize yourself with the raster data model in two ways: First you will convert several data sets to raster data, and in so doing you will become more familiar with some of the differences in how spatial data is looks in the raster data model, both in terms of spatial representation and attribute handling. Second, you will use the Raster Calculator, one of the functions available in the ArcGIS Spatial Analyst Extension, to experiment with map algebra to see how overlay analysis works in the raster data model.

Data

The exercises in the first half of the course will use spatial data from the vicinity of the Coldstream Creek catchment, a small watershed in Okanagan Valley of British Columbia, Canada. This exercise will use the following 5 datasets:

- elevzones50 – This raster GRID divides the catchment into 3 elevation bands
- lulc.shp – This polygon shapefile provides land use/land cover information
- sampling.shp – This point shapefile marks 8 sampling sites and channel characteristics at those locations
- soils.shp – This polygon shapefile provides soil name and texture class information
- streams.shp – This polyline shapefile maps the drainage network

Procedure and Questions

<u>Vector to Raster Conversion</u>

- The *Spatial Analyst* extension must be turned on in ArcMap to provide the capability to work with raster datasets. After starting ArcMap, go to *Tools* → *Extensions* in the menu and make sure there is a check mark beside *Spatial Analyst*. Next, go to *View* → *Toolbars* and ensure there is a check mark beside the *Spatial Analyst* toolbar. You'll need the controls in the *Spatial Analyst* toolbar to perform the operations in this exercise.

- Add the five datasets to your map document

- Before doing anything with the *Spatial Analyst*, it is wise to set the options that will determine the characteristics (extent and cell size) of the GRIDs it will create. Click on the *Spatial Analyst* toolbar's pull-down menu. The menu item at the very bottom should be *Options*. In the *Extent* and *Cell Size* tabs, set the *Analysis extent* and *Analysis cell size* to be *Same as Layer "elevzones50"*. Doing this will cause any

GRID produced by the *Spatial Analyst* in this session to have cells that match perfectly with the cells of the elevzones50 GRID.

- We are now ready to perform some vector to raster conversions. These are accomplished using the *Convert → Features to Raster…* menu item in the *Spatial Analyst* toolbar pull-down menu. The *Input features* item specifies which shapefile you are converting. The *Field* item determines which attribute from the shapefile is used to assign grid cell values in the resulting grid. Use the following fields for each of the four shapefiles you will convert:
    - sampling – SITE
    - streams – L_ORDER
    - lulc – PLU_CLASS
    - soils – TEXTURE

  Leave the *Output cell size* at 50 (for now). You must choose *Output raster* names carefully: GRID datasets are somewhat limited in their naming; use only letters and numbers and do not use any spaces in the file name. As a suggestion, name your GRIDs by appending the cell size to the name (e.g. use the name sampling50 for the GRID you create from sampling.shp; this convention will be used through the remainder of these instructions).Save your GRIDs in your h:\ space (you may wish to create a subdirectory there for each lab exercise).

- As you convert the *Features to Raster*, they will automatically be added to your map document. You may wish to adjust the symbology of these layers as they are added.
- Open the attribute tables of the original shapefiles and those of your new GRIDs.

Question 1 – How do the attribute tables of the original shapefiles and those of the GRIDs you have just created differ? Why is this the case? What could we potentially do about it?

- Have a closer look at sampling.shp and the sampling50 GRID you created from it. You'll need to zoom in on the sampling location points one at a time. Make sure sampling.shp is above your new sampling50 GRID in the list of layers so you can see both once zoomed in.

Question 2 – Which spatial data model more accurately positions the sampling locations? How precisely does your new raster GRID locate them? What property does the raster representation of these points have that the vector representation does not, and is this desirable?

- Now have a closer look at streams.shp and the streams50 GRID you created from it. Again, you may need to zoom in a little to see both clearly. Make sure streams.shp is above your new streams50 GRID in the list of layers so you can see both once zoomed in.

Question 3 – Compare the spatial representations of the streams in the vector and raster data models. What property does the raster representation of these polylines have that the vector representation does not, and is this desirable?

- Compare lulc.shp to the lulc50 GRID you have just created. To do this, draw lulc.shp on top the lulc50 GRID, and set the symbology of lulc.shp to be hollow or somewhat transparent so you can see the GRID below it. Do the same for the soils data.

Question 4 – Compare the spatial representations of the land use land cover and soils polygons in the vector and raster data models. How does the raster version differ from the vector (again, you may need to zoom in a little)?

- Finally, convert each of the 4 shapefiles to raster again, leaving all parameters the same, **EXCEPT** change the cell size to 25. Be sure to name the resulting GRIDs so you can distinguish the versions from one another (e.g. sampling25). Now, compare the 25-meter cell versions to the 50-meter cell versions by placing the vector version on top, the 25-meter version below it, and the 50-meter version on the bottom (in the case of the two polygon layers, you will need to click one of the GRIDs on and off to compare them).

Question 5 – What effect does halving the cell size have on the quality of the raster data model's spatial representation of each of the 4 shapefiles? Comment specifically on how decreasing the cell size changes the representation of point, polyline, and polygon features.

<u>The Raster Calculator</u>

- The *Raster Calculator* function of the *Spatial Analyst* provides a means by which overlay analysis can be performed on GRIDs. By typing an expression into the text entry portion of the *Raster Calculator* dialog, you can combine information from multiple raster GRIDs using a variety of arithmetic and logical operators. Open the *Raster Calculator* by selecting it in the *Spatial Analyst* toolbar's pull-down menu. In the dialog that pops up, you should see all the raster layers you have in your map document listed under *Layers*, along with buttons to help you build an expression by using the graphical user interface (alternatively, you can simply type the expression). You can gain access to more functions (*Arithmetic, Trigonometric, Logarithmic, Powers*) by clicking the >> button.
- As an experiment, let's try creating an expression that will locate all streams within the lowest elevation zone. To do this, we need to first recognize that we must express those two conditions in terms of the values contained in each of these GRIDs. The values of your streams50 GRID reflect the stream order of the pixels (remember, we used the L_ORDER attribute from the streams shapefile to create this grid). Since we are interested in all streams, we want pixels with any stream order, and we can express this using the expression [stream50] > 0. Examining the elevzones50 GRID, we can see that the lowest zone has a value of 1, so we can express this using [elevzones50] == 1 (the double equals is a logical equality). We can combine these two conditions using the and (&) operator: [streams50] > 0 & [elevzones50] == 1. Enter that expression into the *Raster Calculator* and click the *Evaluate* button.

- If the expression was properly formed, the result will be that a new layer named Calculation is added to the map document, with values of 0 and 1. Values of 0 will be present where the expression we evaluated is not true (either where streams50 is not greater than 0 or where elevzones50 is not equal to 1), and values of 1 will be present where the expression is true (where streams50 is greater than 0 and elevzones50 is equal to 1).
- Have a close look at the resulting Calculation layer by clicking off all the checkmarks in the *Table of Contents* so only the Calculation layer is drawn. Note that the resulting layer only has values where there were values in the streams50 layer. This is because in all those locations where streams50 is not equal to either 1, 2, 3, or 4, it has the value No Data. There is a significant difference between the value 0 and the value No Data in a raster GRID. The expression we used earlier cannot be evaluated for any cell which has a No Data value in any of the input layers; this is because the expression is neither true nor false in these locations, it simply cannot be evaluated (thus, a rule to remember is that No Data trumps all … any operation performed on a No Data cell results in a No Data result).
- Suppose we want to identify locations where a stream is not present. To do so, we must reclassify the streams50 GRID, changing the No Data values to 0 values. Open the *Reclassify* dialog by selecting it in the *Spatial Analyst* toolbar's pull-down menu. Use streams50 as the *Input raster* and VALUE as the *Reclass Field*. Place a 0 in the *New values* column beside the No Data value in the *Old Values* column. Leave *Output raster* as *<temporary>* and click the OK button. The result will be that a new layer entitled Reclass of streams50 is added to your map document. This new layer will be identical to the streams50 layer, except the No Data values are replaced by 0 values. We can use this new layer to identify non-stream cells.
- Use the *Raster Calculator* again to find all the streams in the lowest elevation zone, only this time substitute the new Reclass of streams50 for streams50 in the previous expression.

Question 6 – *How does the resulting Calculation2 layer differ from the one we created previously (Calculation) using the streams50 layer as an input?*

- Let's combine several layers and criteria now. Suppose we want to find cells at lower elevations that might be suitable for conversion to crop agriculture: Create an expression in the *Raster Calculator* that finds cells that:
    - Do not contain streams
    - Are in the lowest elevation zone
    - Are currently classified as rangelands or youngForest in the lulc50 GRID
    - Have soil type SL (sandy loam)

Question 7 – *What is the proper expression to find the cells with the above criteria (**Hint: You'll need to use some brackets**)? Print out a map of your resulting Calculation GRID, and title the map with the expression you used to create it.*