Exercise 10: Calculating Sigma from LST-NDVI Data

Introduction

Now that we have become familiar with the operations required to convert remotely-sensed reflectance and brightness temperature data into normalized difference vegetation index (NDVI) and land surface temperature (LST) respectively, we can begin to make use of these data to derive a description of surface moisture conditions.

The theory we are relying upon here requires that we build a 2-dimensional distribution from the surface temperature and vegetation index information. The T_s-VI distribution places the vegetation index on the x-axis and the surface temperature on the y-axis. For an identified region (i.e. the set of pixels of interest) the T_s and VI values for each pixel are collected and plotted. We can then interpret the shape of the distribution that is formed. One interpretation considers the T_s-VI distribution to be triangular, and the upper edge of the triangle is the dry line. The slope of the dry line can be used as an indicator of the overall moisture condition of the region (Nemani and Running. 1989; Nemani et al., 1993). In this exercise, we will collect the T_s-VI distribution from a remotely-sensed image, and fit a line to the upper edge to define the dry line, and find the slope of that line.

Data

The exercises in the remainder of the course will use remotely-sensed data in the vicinity of North Carolina Climate Division 3. This geographic unit is an agglomeration of 13 counties in north-central North Carolina, which includes the cities and towns of Durham, Chapel Hill, Greensboro, Burlington, and Winston-Salem. This particular exercise will make use of land surface temperature (LST) and normalized difference vegetation index (NDVI) datasets collected by the MODIS Terra sensor on May 24, 2002:

- ids – This shapefile contains square shapes for each of the pixels in the remotely-sensed imagery of North Carolina Climate Division 3
- lst – This GRID contains land surface temperature values in degrees Celcius
- ndvi – This GRID contains normalized difference vegetation index values

Procedure and Questions

<u>Forming the T_s-VI Distribution</u>

- While the ArcGIS *Spatial Analyst* is a versatile and powerful tool for working with raster datasets, it does have some limitations. One thing that it is not designed to do is to manipulate/work with the values of individual pixels (to put this another way, the *Spatial Analyst* functions are generally designed to approach a whole GRID as a unit and to operate on GRIDs as a whole).
- The procedure we will undertake in this exercise in forming the T_s-VI distribution and finding the slope of the dry line will require us to work with the values of

individual pixels quite explicitly. For this reason, we will need to make use of some tool sets and capabilities other than the *Spatial Analyst* to accomplish this task.

- The first capability we must take advantage of is ArcGIS' integration of the vector and raster data models. While we cannot work with the values of individual pixels using *Spatial Analyst*, we can make use of a specially-designed shapefile to collect the values of individual pixels in a table. The provided shapefile ids.shp is such a shapefile: It contains square polygons that each have the same extent as one of the pixels in the North Carolina Climate Division 3 imagery (there are a total of 14931 of these).
- We will use *Spatial Analyst's Zonal Statistics* function (found in the *Spatial Analyst* menu) to collect the required values. The *Zone dataset* will be ids shapefile, the *Zone field* will be <Value>, the *Value raster* will be the ndvi GRID (at first, we will do this again with the lst GRID), leave *Ignore NoData in Calculations* checked, uncheck *Chart statistic*, and choose an *Output table* location in your H:\ space. Once you click *OK*, the table will be created, opened in ArcGIS, and saved as a .dbf where you specified.
- Repeat the above procedure for the lst GRID. You should now have two new .dbf files at specified locations in your H:\ space. They each will contain 14930 records (note, the reason this is not 14931 is because there is a single *NoData* pixel in the imagery that produces no record in the tables). We are now ready to begin working with these two tables in Microsoft Excel.
- Open the NDVI table in Excel. Delete all the fields except FID_ and MEAN. Rename the MEAN field to NDVI. Reduce the number of decimal places shown for the NDVI field to 4.
- Open the LST table in Excel. Copy the MEAN field to the NDVI table. Once there, rename it to LST, and reduce the number of decimal places shown for the LST field to 2.
- You should now have a worksheet that contains 3 fields: Column A should contain FID_, Column B should contain NDVI, and Column C should contain LST. Save this as an .xls file in your H:\ space.
- You now should have the data in a form where it will be easy to create the T_S-VI distribution scatterplot. Select the B and C columns and use the *Chart Wizard* to create a scatterplot of the values. Use the *XY (Scatter)* chart type, and the *Scatter* sub-type where the points are **NOT** connected.

Question 1 – *Is the T_S-VI distribution reasonably triangular in form? To supplement your descriptive answer, create and hand in a plot of the T_S-VI distribution (use a scatterplot as described above, make sure the NDVI is on the x-axis, the LST values on the y-axis, with both axes properly labelled and scaled to the data range {right click on the axes and use Format Axis, setting the NDVI range to be 0 – 1 and the LST range to be 20 - 45}, and that the plot includes a title that is descriptive and appropriate).*

Question 2 – *Could you fit a line to the dry edge from the complete distribution of T_S-VI values from the image? Would this be easy or difficult to do? What difficulties would you run into and how might you solve them?*

Finding the Upper Envelope of the T_S-VI Distribution

- In order to make use of an automated line fitting procedure to find the dry edge of the T_s-VI distribution, we need a means of selecting just those pixels/points that form the upper envelope of the distribution. There are any number of possible approaches that we can take to accomplish this task, but we shall make use of one that requires no creation of custom code, but rather uses the capabilities built into ArcGIS and Excel.
- What we are essentially looking for is the points with the highest LST values for a given (range of) NDVI value(s). We can find these through an approach that reclassifies NDVI into a set of intervals. We can then use the resulting classes as zones in the *Spatial Analyst's Zonal Statistics* function, and use the Maximum value of LST from each NDVI zone to represent the highest LST within the desired NDVI range (which, in turn, can be represented by the mean NDVI of that range).
- To implement this, begin by reclassifying the ndvi GRID. Use the *Spatial Analyst Reclassify* function (found in the *Spatial Analyst* menu). Select ndvi as the *Input raster*, <Value> as the *Reclass field*, and click the *Classify* button: Select the *Equal Interval* method with 32 classes, and click *OK*. Select a location for the *Output raster* in your H:\ space and click *OK*. This will create your 32 NDVI zones GRID.
- Now, use the *Spatial Analyst's Zonal Statistics* function as you did before, but this time with your reclassified NDVI grid as the *Zone dataset*, on both the NDVI and LST data. Save the output tables to appropriate locations in your H:\ space.
- Open these two new .dbfs in Excel (they should each contain 32 records) to assemble your new T_s-VI distribution consisting solely of the upper envelope. Remember to use the MEAN value from the NDVI table you just created, and the MAX value from the LST table you just created (remember to *Format Cells* appropriately, with 4 decimal places for the NDVI data, and 2 for the LST data). Again, you should end up with a table containing 3 columns: Column A should contain VALUE (indicating the 32 NDVI intervals), Column B should contain NDVI (populated by the MEAN NDVI value from the intervals), and Column C should contains LST (populated by the MAX LST value from the intervals). Save this as an .xls file in your H:\ space.
- Once again, create a scatterplot as you did last time.

Question 3 – *Does this look like a set of points to which we can easily fit a line? To supplement your descriptive answer, create and hand in a plot of the upper envelope of the T_s-VI distribution (use a scatterplot as described above, make sure the NDVI is on the x-axis, the LST values on the y-axis, with both axes properly labelled and scaled to the data range {right click on the axes and use Format Axis, setting the NDVI range to be 0 – 1 and the LST range to be 20 - 45}, and that the plot includes a title that is descriptive and appropriate).*

Question 4 – *If we are going to make use of an automated slope fitting procedure, what problems might we have with this dataset of 32 points in its current form, and what can we do about these problems?*

Fitting the Dry Line Slope to Obtain σ

- Looking at the resulting 32 points in our derived upper envelope distribution (from our selected method, which is admittedly a little less sophisticated than operational methods that have been described in the readings, which start with all the values, and iteratively remove them) we can see that a few points in the lower NDVI intervals appear to not be part of the true upper envelope. By eliminating these points from consideration, we can safely use least squares regression to fit a line and find its slope.
- Within your workbook, make a copy of the worksheet containing the 32 records. Next, in the copy, eliminate any rows that contain an LST of less than 30 degrees Celsius.
- We can now use one of Excel's built-in functions to find the slope of a regression line fitted to these points. Choose an empty cell in your newly-created worksheet, and click on the *f(x)* button. In the *Statistical* function category, you will find the *Slope* function. Fill in the required parameters (the LST values are the *known_y's* and the NDVI values are the *known_x's*) and click *OK* to calculate the resulting slope.

Question 5 – What is the slope of the linear regression line fitted to these points?

- Again, make a scatterplot of the now pruned upper envelope of the T_s-VI distribution (as for the previous two, with all the same characteristics). Right click on any data point in the plot and select *Add Trendline…* to the plot, using a *Linear* trendline. This adds the resulting regression line to the plot.

*Question 6 – Does the linear regression fitted trendline do a good job of capturing the slope of the upper envelope of this T_s-VI distribution? To supplement your descriptive answer, create and hand in a plot of the pruned upper envelope of the T_s-VI distribution (use a scatterplot as described above {**this one should include the trendline**}, make sure the NDVI is on the x-axis, the LST values on the y-axis, with both axes properly labelled and scaled to the data range {right click on the axes and use Format Axis, setting the NDVI range to be 0 – 1 and the LST range to be 20 - 45}, and that the plot includes a title that is descriptive and appropriate).*

*Question 7 – Perform this procedure using only intervals with mean NDVI values greater than 0.5 (this is an accurate reflection of where the majority of the original 14930 points are located) by copying the remaining values to another worksheet and eliminating some as described. What is the resulting slope in this case? Again, create a plot with a trendline, and comment on the quality of fit in this case as compared to that in the plot for Question 6. To supplement your descriptive answer, create and hand in a plot of the further pruned upper envelope of the T_s-VI distribution (use a scatterplot as described above {**again, include the trendline**}, make sure the NDVI is on the x-axis, the LST values on the y-axis, with both axes properly labelled and scaled to the data range {right click on the axes and use Format Axis, setting the NDVI range to be 0 – 1 and the LST range to be 20 - 45}, and that the plot includes a title that is descriptive and appropriate).*