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 11: Calculating TVDI from LST-NDVI Data

### Introduction

In our previous exercise, we built a 2-dimensional distribution from surface temperature and vegetation index information derived from remotely-sensed imagery. 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 then interpreted the shape of the distribution that is formed as triangular, and found the slope of the upper edge of the triangle (the dry line) to provide an indicator of the overall moisture condition of the region (Nemani and Running. 1989; Nemani et al., 1993).

In this exercise, we will go one step further, and derive a per-pixel description of moisture condition that can be mapped back to geographic (as opposed to parameter) space. This procedure still begins with collecting the $T_s$-VI distribution from a remotely-sensed image, and fitting a line to the upper edge to define the dry line, and finding the slope of that line. However, it proceeds to also define the lower edge of the triangle as the wet line. Together these two lines form a description of the range of wetness/surface temperatures at a given VI value. We can then apply the temperature vegetation dryness index (TVDI), which operationally is an algebraic approach to calculating a given pixel's position in the $T_s$-VI distribution with respect to the dry and wet lines (Sandholt et al., 2002; Andersen et al., 2002).

### 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

Forming the $T_s$-VI Distribution

- As was the case in the previous exercise, we will need to make use of a combination of capabilities of ArcGIS and Excel to accomplish the task. We will once again need to manipulate the values of individual pixels.
- We will begin once again by forming the $T_s$-VI distribution. If you have saved an Excel file from the previous exercise containing the 14930 identifiers, NDVIs, and LSTs from this set of imagery (the same as in Exercise 10), you may make use of that file and skip the next few steps that are needed to recreate it (move on to the next section/page). Otherwise, start ArcMap and add the three datasets described above to your map and proceed as described below.
- 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 has 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 VALUE 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 VALUE, Column B should contain NDVI, and Column C should contain LST. Save this as an .xls file in your H:\ space.

Defining the Edges of the $T_s$-VI Distribution – The Dry Line

- 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.
- Recall that in Exercise 10, we derived two lines fitted to the upper envelope of the $T_s$-VI distribution, and these had the following equations:

  y = -10.966x + 45.235 (dryline1)

  y = -21.441x + 53.181 (dryline2)

- Both of these line fits were derived through a simplified method. A further dry line for this dataset, derived through a more involved method is:

  y = -20.7001x + 50.3325 (dryline3)

- For our next step, we will plot each of these three dry lines on top of the full $T_s$-VI distribution. Begin by sorting the three columns (A through C) by ascending NDVI (be sure to use the *My list has Header row* option). Once you do so, the first record should be VALUE 9233 with an NDVI of 0.0644 and the last record should be VALUE10644 with an NDVI of 0.9546.
- Add 3 new column headers in cells D1, E1, and F1: Label them dryline1, dryline2, and dryline 3 respectively.
- Calculate the modeled minimum and maximum LST values (across our NDVI range) for each of our 3 dry lines, using the aforementioned minimum and maximum NDVI values and the equations above (e.g. in cell D2, calculate the LST for the lowest NDVI in the image using the dryline1 equation, such that the value in cell D2 = (-10.966 * cell B2) + 45.235, repeating this for cells E2 and F2 with the other equations, and cells D14931 through F14931 for the maximum values).
- Add these data to the chart using the *Add Data* menu item when the chart worksheet is selected. Use all of columns D through F as the *Range*, *Add Cells as New Series, Values (Y) in Columns,* and *Series Names in First Row*.
- Add a trendline for each of our three new data series. Once you have created the trendlines, right click on the lines themselves to *Format Trendline* and modify their symbolizations so the three can be distinguished from one another.

*Question 1 –* *What are the modeled minimum and maximum LST values (across our NDVI range) associated with each of the three dry lines specified above?*

*Question 2 –* *Discuss how effectively the three dry lines fit the upper envelope of the $T_s$-VI distribution. To supplement your description, print out and hand in the plot of the $T_s$-VI distribution that has the dry lines plotted (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 - **55**}, make sure **the legend is shown** to allow the various series to be distinguished from one another, and that the plot includes a title that is descriptive and appropriate).*

*Question 3 – In all three cases, some points appear above the fitted dry lines. What would this mean for the TVDI calculated for those pixels if we used the respective dry lines for the calculation (refer to Sandholt et al., 2002 or the lecture materials to refresh your memory of your TVDI works)? How might we deal with this difficulty?*

Defining the Edges of the $T_S$-VI Distribution – The Wet Line

- The wet line is a little more straightforward to define than the dry line, in that the convention is to make use of a line with zero slope at a selected land surface temperature (that is, we usually parameterize the triangle to have a flat bottom, although we conceivably could do so another way).
- The simplest approach is to set the wet line to be equal to the lowest temperature observed in the dataset. We can use similar techniques as employed above to find where that line should be located and display it on another $T_S$-VI distribution plot.
- Begin by copying the worksheet with the data values, and removing the dry line information from columns D through F of the new copy.
- Add 2 new column headers in cells D1 and E1: Label them wetline1, and wetline2 respectively.
- Now, sort the columns by ascending LST to find the lowest surface temperature. Make a note of it, and then re-sort the columns by ascending NDVI (in both cases, use the *My list has Header row* option).
- Fill in the values for cells D2 and D14931 (these should be the rows with the min. and max. NDVI values) with the lowest surface temperature you took a note of previously.
- An alternative way of defining the wet line is to attempt to minimize the effect of noisy points that sit at the bottom of the distribution. That is, we can discard from consideration the lowest 5% of the temperature observations in defining the wet line.
- To do so, once again sort the data by ascending LST (make sure you select all 5 columns before sorting, and use the *My list has Header row* option). Given that there are 14930 observations, we can take the LST value in row 748 to be the 5th percentile of LST observations (14930 * 0.05, rounded up, plus 1 since the first observation is in the 2nd row). Make a note of the LST in row 748, then select the 5 columns and resort them back to ascending NDVI (as always, using the *My list has Header row* option).
- Enter the appropriate LST values in cells D2, E2 (the minimum LST from the NC CD3 pixels), D14931, and E14931 (the 5th percentile LST from the NC CD3 pixels) to define our two wet lines. Then, use the same series of operations as you did above for the dry lines to create a plot that includes the $T_S$-VI distribution and the two wet lines.

*Question 4 – What are the minimum and 5th percentile LST values from NC CD3 for our image?*

*Question 5 – Discuss how effectively the two wet lines fit the bottom of the $T_s$-VI distribution. To supplement your description, print out and hand in the plot of the $T_s$-VI distribution that has the wet lines plotted (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}, make sure* ***the legend is shown*** *to allow the various series to be distinguished from one another, and that the plot includes a title that is descriptive and appropriate).*

*Question 6 – In the case of one of the wet lines, some points appear below the wet line. What would this mean for the TVDI calculated for those pixels if we used that wet line for the calculation (refer to Sandholt et al., 2002 or the lecture materials to refresh your memory of your TVDI works)? How might we deal with this difficulty?*

## Calculating TVDI and Mapping it Back to Geographic Space

- We will use the equation defining dryline3, and the 5th percentile LST defining wetline2 for our TVDI calculation.
- Recall the form of the TVDI equation:

$$\text{TVDI} = (T_S - T_{SMin}) / (a + b\text{VI} - T_{SMin})$$

where:

$T_S$ is the LST for the pixel

$T_{SMin}$ is the wet line level (for us, the 5th percentile LST)

a is the dry line intercept

b is the dry line slope

VI is the NDVI for the pixel

- Make another copy of the worksheet containing the data, such that it initially contains just the 3 original columns (VALUE, NDVI, and LST).
- Fill in cell D1 with a new column header, TVDI.
- Translate the TVDI equation shown above to an appropriate expression in Excel, and use it in cell D2 to calculate the TVDI for that pixel (if you did this correctly, given NDVI = 0.0644 and LST = 25.67 you should have calculated TVDI = -0.1254, rounded to 4 decimal places).
- Copy the expression to cells D3 through D14931 to calculate TVDI values for the remaining pixels, and save your Excel workbook.
- Now, use the *Save As* feature to save the worksheet containing the TVDI values to a comma-delimited (CSV) text file (this a file format that ArcGIS can read easily) in your H:\ space.
- Open ArcMap and add the three dataset for this exercise to the map document (if you have not already done so).

- Add the CSV text file to the map document using *Add Data*.
- At this point, we will perform a relational join between the ids.shp shapefile and the CSV file containing the TVDI values. Inside the *Data Management Tools* toolbox, you should find *Joins*, and within that sub-section, *Add Join*. The *Layer Name* will be ids.shp, the *Input Join Field* will be the unique identifier, the *Join Table* will be the CSV file, and the *Output Join Field* will again be the identifier (VALUE).
- Once you have completed the join, open the attribute table for ids.shp. You should now find the columns from the CSV file have been joined to it, including the TVDI values.
- In the *Spatial Analyst* pull-down menu, under *Options*, use the *Extent* and *Cell Size* tabs to set the analysis characteristics to be equal to either the NDVI or LST GRID.
- In the *Spatial Analyst* pull-down menu, use *Convert*, *Feature to Raster* to create a TVDI GRID. The *Input features* are id.shp, the *Field* is TVDI, the *Output cell size* should remain at 1000, and select an appropriate *Output raster* location in your H:\ space.

*Question 7 – Compare your resulting TVDI GRID to the map of North Carolina Climate Division 3 land-use land-cover that can be found in the accompanying lecture PDF. Are there any common features that can be found? What are they and why would you expect them to be expressed in a GRID of TVDI? Include a printed map of the TVDI inside NC CD3 GRID with your exercise, and make sure it has the following characteristics: It should have an appropriate title, north arrow, scale, and legend (be sure to use the Layout View for this).*