Geography 378: Introduction to Geocomputing
Lab 5: GDAL and Python

Assigned: 11/10
Due: 11/16
15 points

Hand-in

- Please collect your answers in a single .py file. Compress your Python file and the image you generate into a single .zip file and name your zip file as **lab5_yourname.zip.**
- Submit the zip file to the assignment folder called "Lab 5".
- Include appropriate comments to explain what each line or block of code accomplishes. **You must comment your code for full credit.**

Getting Started

TIFF files (.tif or .tiff) are a raster format commonly used for GIS because they can store spatial reference information in their headers and are read by most image manipulation and GIS software. If you examine a typical RGB tiff file using GDALINFO, you will see that it consists of three bands: Red, Green, and Blue. When you Preview the image or open it in an image manipulation program (e.g. Photoshop), the program automatically assigns each band to one of the three color channels in your computer monitor; the colored light produced by each channel blends together to create the colors you see on the screen.

Some raster files have more or fewer than three bands. The minimum number of bands in a raster is one. Any less and the file would be an empty container. Since rasters often represent useful containers for spatial data it is possible for rasters to have many thousands of bands. In spatial Bayesian analysis each band could represent a posterior draw, so you could have thousands of representations of the same spatial process.

The files in this Lab are all single band rasters. For this lab there are 6 files, each of which comes from a Landsat 7 satellite image of the Mississippi River. The pixel values represent the amount of electromagnetic radiation at various wavelengths captured by the satellite for that pixel area. In this case you can think of the collection of files as a multi-band image of the area, with each file storing a different "color" or wavelength. In this image set, band 1 is blue (i.e., file 1 has blue radiation amounts), band 2 is green, band 3 is red, band 4 is near infrared, and bands 5 and 6 are different areas of shortwave infrared light.

One common use for Landsat data is to analyze the biological productivity of an area, as determined by the amount of vegetative cover. A relatively simple index, NDVI (Normalized Difference Vegetation Index) is often used to show areas of high and low vegetative growth. The NDVI calculation produces a gray-scale raster image with pixel values ranging from -1 to 1; darker pixels (i.e. values between -1 and 0) show areas of less vegetation, while lighter areas (i.e. values between 0 and 1) have more lush vegetation. The NDVI equation is:

$$\text{NDVI} = \frac{\text{Near IR} - \text{Red}}{\text{Near IR} + \text{Red}}$$

Lab Task

Using GDAL, write a Python script to produce an NDVI image from the provided Landsat data.
Directions:

- The Near IR band of the image is band 4 (_B40_ in the filename), and Red is Band 3 (_B30_).
- Be careful with issues that may arise from dividing by zero when operating on the raster arrays. Hint: To avoid it, add ".astype(float)" after "ReadAsArray()" method.
- Close the file when finished.
- Your script should not crash in any case. If the file does not exist, you need to print a warning message and exit. Hint: If a file cannot be opened via gdal.Open() method, the value of the file is "None".
- The final image created by the script should be readable in ArcGIS or QGIS.
- Submit the image in TIF/TIFF format together with your script.