Summer Institute in Earth Sciences 2010

Assessing the Accuracy of Land Cover Estimates for Earth Satellite Applications

Ryan Max
University of Maryland Baltimore County

Mentor: Dr. Eric Brown de Colstoun
University of Maryland Baltimore County
Goddard Earth Sciences and Technology Center
NASA Goddard Space Flight Center

Introduction:

The work that I've done in this internship has been about evenly divided between three separate projects, all dealing with Landsat land cover data in some form. The first project was an evaluation of the precision of field-based measurements of tree cover, also called "ground-truthing." This project required only a general familiarity with statistics and to become acquainted with a measurement protocol. The second project was the creation of land-cover maps of Goddard Space Flight Center and the surrounding areas using unsupervised classifications. The third project was the creation of pansharpened classifications of urban subsets around the eastern corridor of the United States into impervious and nonimpervious classes for use as training data for Landsat. The latter two projects required that I learn about the fundamentals of remote sensing, how to process multispectral and multitemporal data and how to run classifications using Envi software. The acquisition of these skills required the plurality of my time as an intern at Goddard this summer. Because there has been no single research project to which I have made large contributions, I will instead detail the work that I have done in each of these areas, and in the case of the first project, some preliminary conclusions.

Methods:

For the field measurement project, we used a field protocol that borrowed heavily from the Globe Program's sample site and biometry protocols.

Setting up the site

After selecting a homogenous site, a handheld GPS unit is used to locate the center of the Landsat pixel associated with that site, which is marked with a flag for future reference. Using a compass and tape measure, the corners of the pixel are marked 21.21 meters to the NW, NE, SW, and SE. In order to compensate for error in the GPS unit and the georeferencing of the Landsat pixel itself, we extended the site measurement by another 15 meters in each direction from the usual $30m^2$ pixel site. Therefore, each transect is extended another 21.21 meters from each corner, and the total distance across the diagonal transect of the site is 84.84 meters. A diagram of the site can be seen in Figure 1. The corners of both the actual pixel and the total site are marked with flags for use during the measurement.

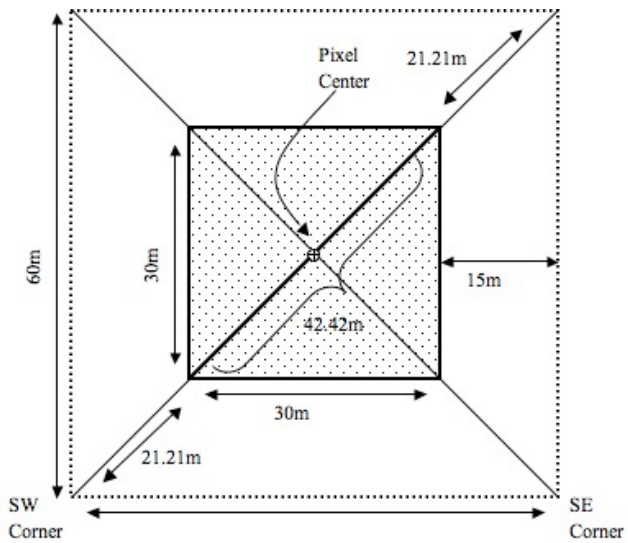

Figure 1: Diagram of sample site, 60m x 60m

<u>Taking the measurement:</u>

In taking the tree cover measurements from the sample site, we followed the Globe Program's biometry protocol that is provided for teachers to allow them to take measurements with their students and contribute to the Globe dataset. Because the Globe Program is designed to be accessible to the general public, the protocols are fairly simple and the instruments used are inexpensive.

Each measurement consisted of walking the length of each transect of the sample site and making a tree cover observation every two steps. The observation consisted of looking straight up through a very simple densiometer consisting of a tube, a washer, tape, and string, assembled as seen in Figure 2.

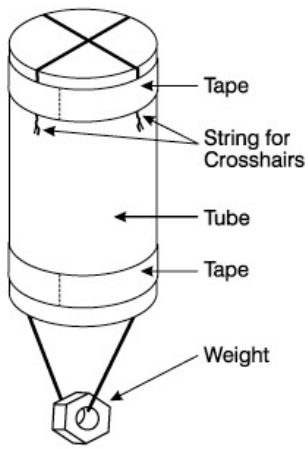

Figure 2: The densiometer used in the Globe Program

If the center of the crosshairs fell on a leaf or branch of the canopy, a positive result was recorded for tree cover, as well as the type of canopy cover (deciduous or coniferous). If the crosshairs fell on a gap in the canopy, a negative result was recorded. Starting from a corner of the pixel, these measurements were taken every two steps, and the beginning and ending of the "actual pixel" region of the site was noted for data analysis purposes. Ground cover data was recorded at each pace as well, but this data was not included in the analysis of the site.

Data Analysis and Conclusions:

The data collected in the field was converted into percent tree cover by dividing the number of positive recorded observations by the total number of observations taken. The percent tree cover measurements for both the pixel and total site are displayed in Figure 3, below, along with trendlines to show the approximate average value over time.

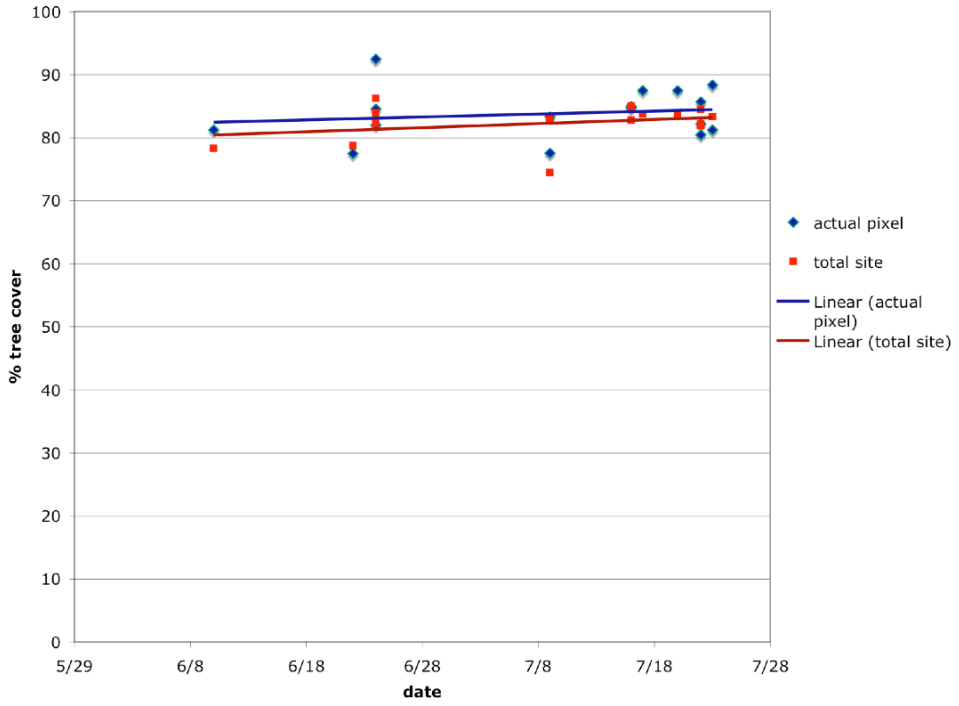

Figure 3: Graph of % tree cover values from each measurement and the date of acquisition.

There are 16 different measurements, with values for both the actual pixel and the total site. The accuracy of these measurements is difficult to ascertain, because the field measurements or "ground-truth" are in fact the measurement against which satellite estimates are validated for accuracy. However, the standard deviation of the measurements gives an idea of the precision of the measurement, with the standard deviation of the pixel measurements at 4% and that of the total site at 2.9%.

One source of error associated with these measurements is the sampling error from pacing only the diagonal transects of the site. Because only the transects are measured, any feature present in the transect that is not representative of the overall tree cover of the site will create a deviation from the percent tree cover of the site. For instance, Figure 4 shows the measurements from each of the total transects separately, and in the case of our site a downed tree meant that the SE to NW transect had an average of 10% less tree cover than the SW to NE transect.

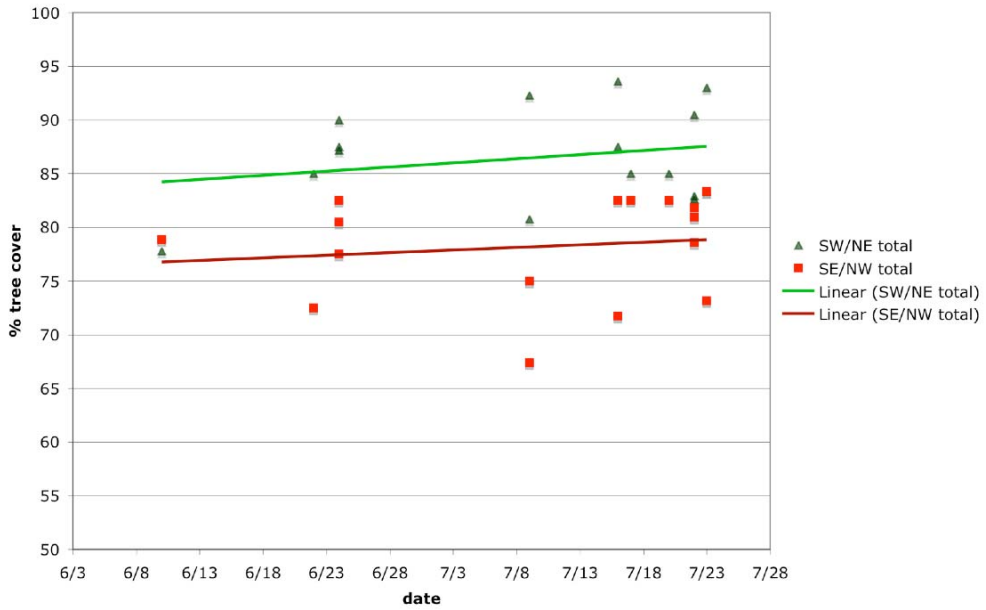

Figure 4: Graph of % tree cover values for each transect and the date of acquisition.

Another source of error in the field measurement protocols is that each person measuring will have slightly different measurements. Different tree cover values can be reached because the measurement is subjective (what constitutes the canopy, whether a leaf blown into the center of the densiometer should be a positive measurement, etc.), because of different stride lengths, and the way in which the densiometer is held. Additionally, when one person makes repeated measurements of a particular site, their precision increases. In our case, I personally made 10 repeat measurements of the same site. In comparing the first five measurements I made with the second five, I found that the range and standard deviation of my measurement values decreased dramatically in both the pixel and total site measurements, as shown below in Figures 5 and 6. This could be partially due to an increased familiarity with the process and a more consistent procedure in taking the measurements, and partially due to an unconscious bias towards a particular value after becoming familiar with the site and processing data on it.

Ultimately, while it appears that the field-based estimates are reasonably accurate, it would require further testing to evaluate exactly what the precision of these measures are, and what conditions and protocols facilitate uniform measurements.

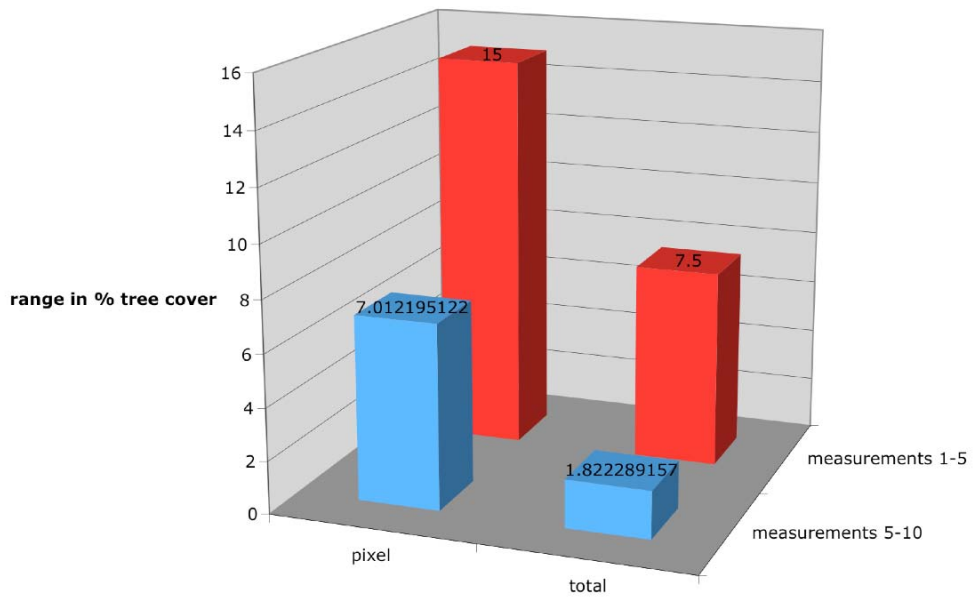

Figure 5: Graph of the data range in % tree cover for both the actual pixel and total site for the first five measurements taken compared to the second five measurements.

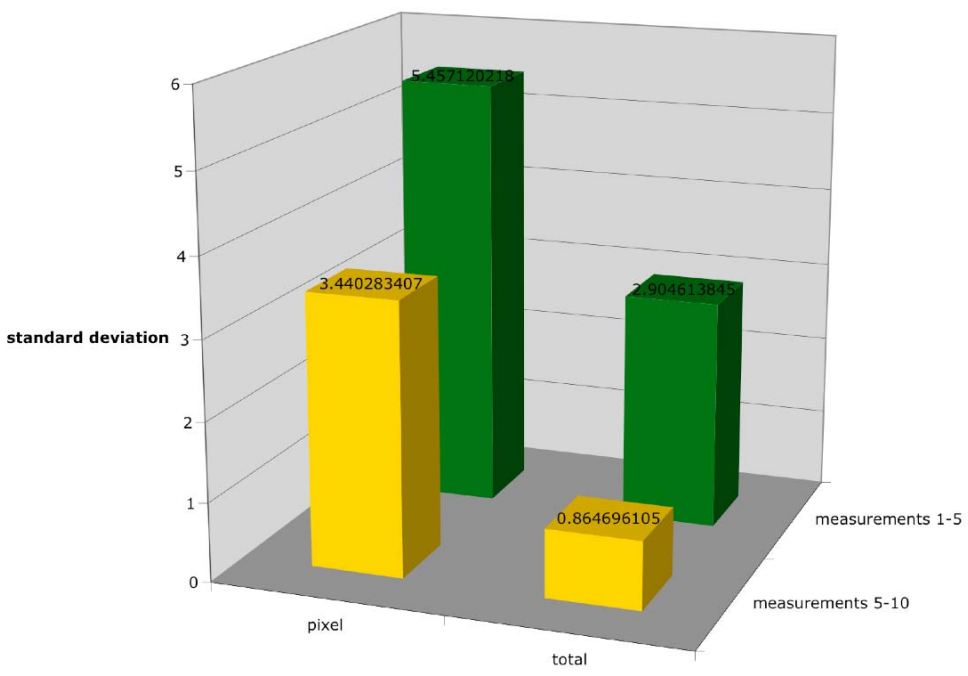

Figure 6: Graph of the standard deviation in % tree cover for both the actual pixel and total site for the first five measurements taken compared to the second five measurements.

Remote Sensing and Land Cover Mapping:

A large portion of my ten-week internship here was spent learning the basics of remote sensing from my mentor. I learned about active and passive sensors, specifically as it pertained to Landsat TM's multispectral data, signature reflectances, and false color composite images. We covered both Multispec, a free piece of software downloadable from the internet, and Envi, a piece of commercial software that we used for unsupervised Isodata clustering classification of the Baltimore-Washington Metropolitan Area. I also learned the importance of small and large scale land cover maps for various interests such as land use/management and monitoring changes in land cover for environmental models. Below are a series of multispectral images of the area between Baltimore and Washington D.C., including GSFC, in "normal color" (Figure 7), "color infrared" (Figure 8), and "false color" (Figure 9) (Lillesand and Kiefer, 1987)

Figure 7: Landsat image of the subset between Baltimore and Washington D.C. This image is in "normal color", with band 3 assigned to red, band 2 assigned to green, and band 1 assigned to blue.

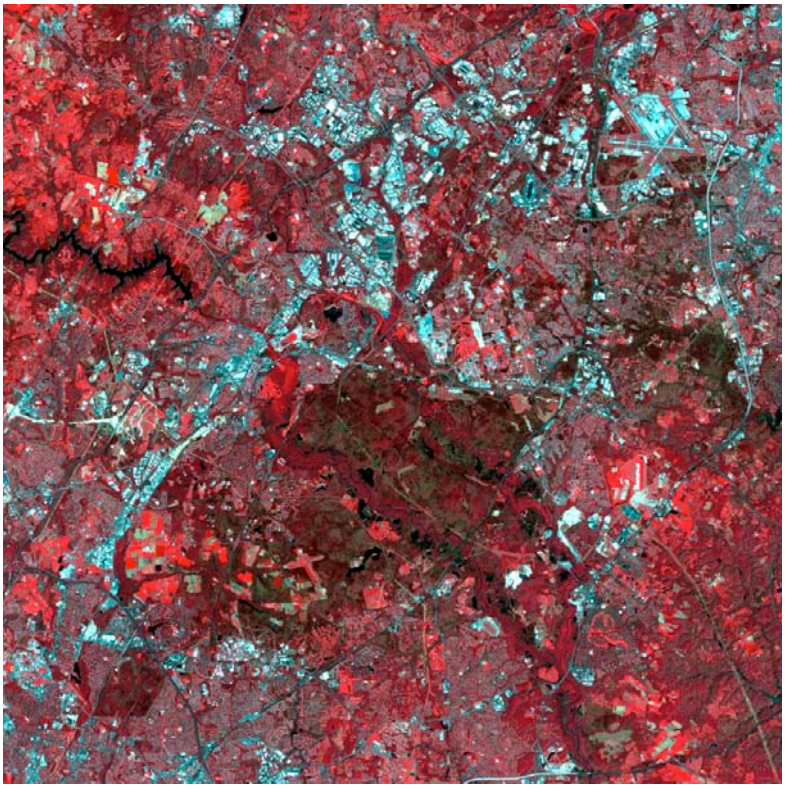

Figure 8: Landsat image of the subset between Baltimore and Washington D.C. This image is in “color infrared”, with band 4 assigned to red, band 3 assigned to green, and band 2 assigned to blue.

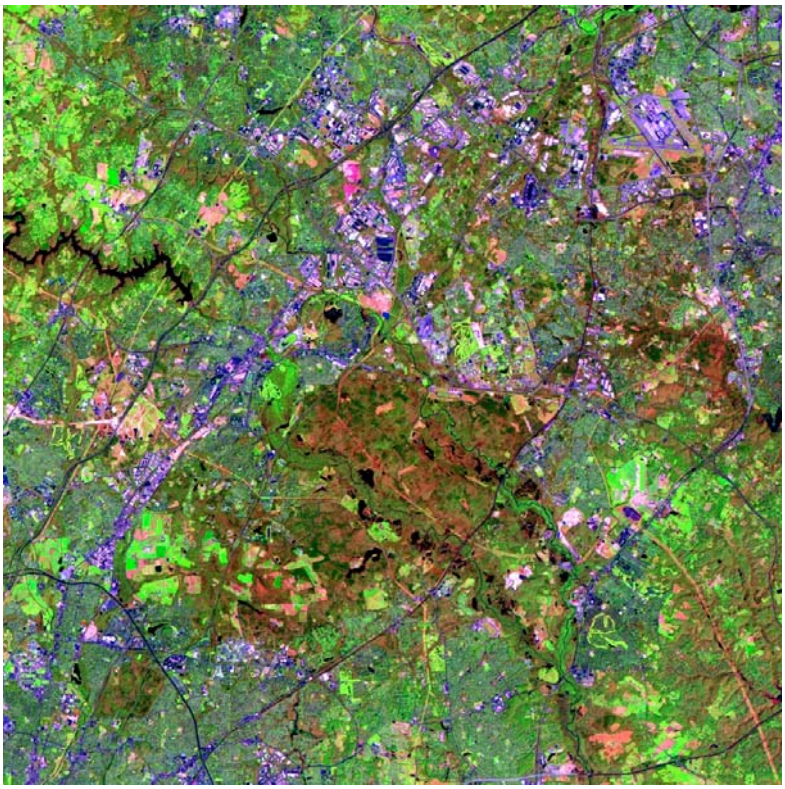

Figure 9: Landsat image of the subset between Baltimore and Washington D.C. This image is in “false color”, with band 5 assigned to red, band 4 assigned to green, and band 2 assigned to blue.

We used the six spectral bands of Landsat ranging from blue light to mid infrared in an Isodata clustering algorithm to create unidentified spectral classes then used various sources to classify each one (Lillesand and Kiefer, 1987). Each of the composites shown above proved useful in identifying different features and types of land cover during the unsupervised classifications of this area. I also used high-resolution satellite data and aerial photography from maps.google.com and from bing.com as tools for validating the Landsat image classifications. Land cover classifications are useful as a visual representation of a large amount of spectral data. By running a classification, the combined data of 6 digital numbers each for thousands of pixels are converted into an image that is visually comprehensible and meaningful, even to people without a scientific background.

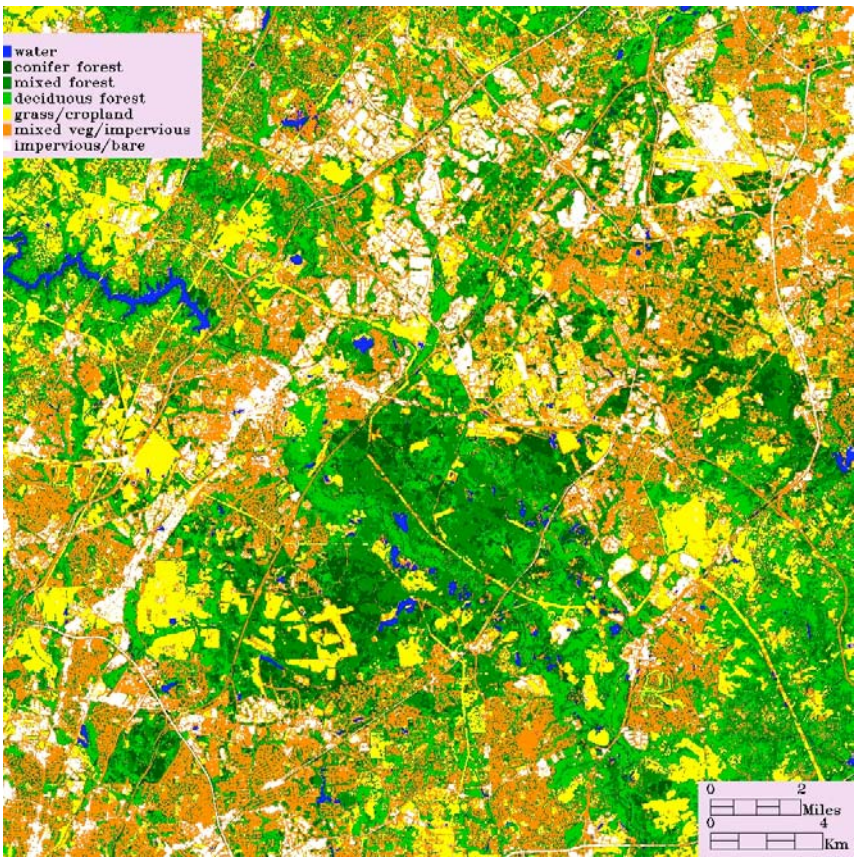

Figure 10: Land cover classification of the subset between Baltimore and Washington D.C. Goddard Space Flight Center is in the SW corner and BWI Airport is in the NE corner.

High-Resolution Urbanization Training Data:

In order to help create a global scale impervious cover map using Landsat data, we contributed higher resolution images classified into impervious and non-impervious cover. In order to create higher resolution classifications, we used IKONOS data, which has spectral data at a spatial resolution of 4m and panchromatic data at a resolution of 1m (Hansen et al, 2002). In order to have a color image at a spatial resolution of 1m to use for classification, we used a technique Gram-Schmidt spectral sharpening (Shah et al, 2008). The result is seen below in Figure 11.

Figure 11: Pan-sharpened IKONOS image of the Capitol Building in Washington D.C. This image is in "true color", with band 3 assigned to red, band 2 assigned to green, and band 1 assigned to blue.

After the pan-sharpened image is produced, we classified various subsets of urban areas on the eastern seaboard into impervious and non-impervious classes. Since the resolution on the classification was at a spatial resolution of 1m, the resulting subsets could be aggregated into 30m pixels and used as training data for Landsat images (Hansen et al, 2002).

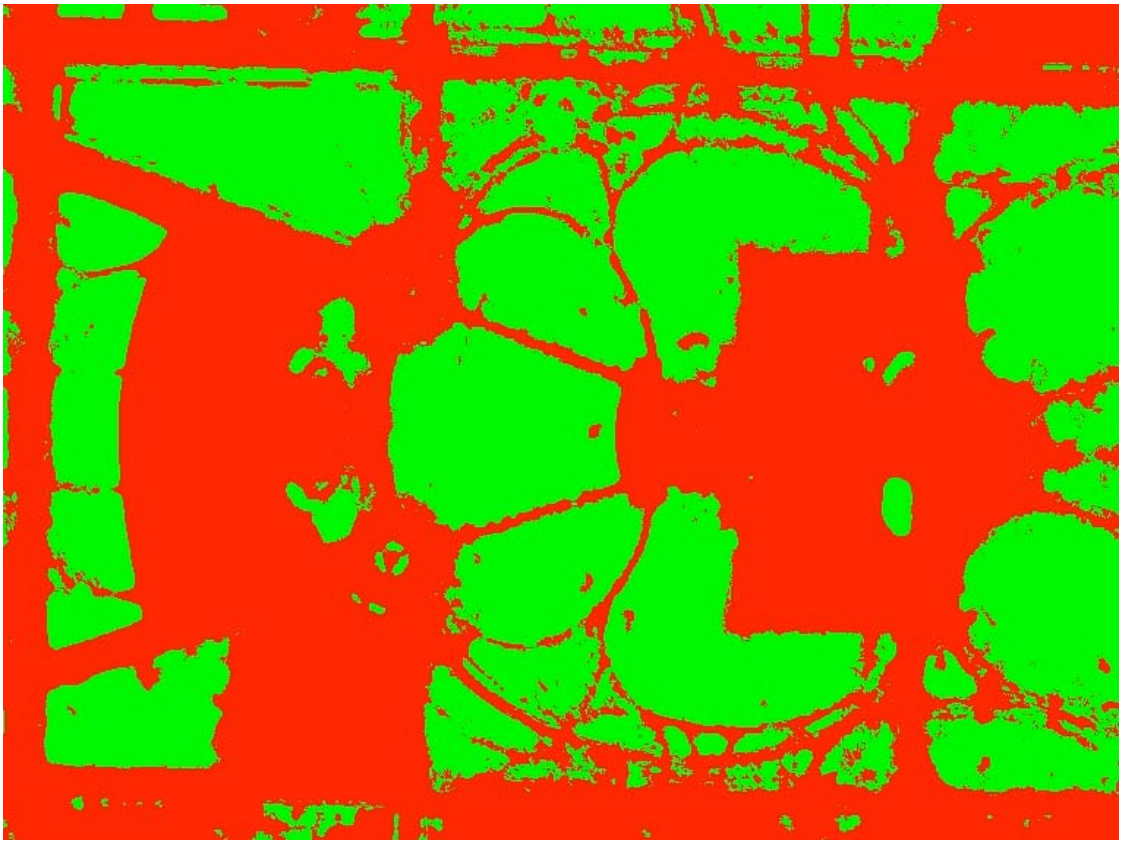

Figure 12: Classification of the IKONOS image of the Capitol Building in Washington D.C. Impervious cover is displayed in red, while nonimpervious cover is displayed in green.

Urban sprawl has changed significantly over the course of the Landsat Data Continuity Mission, and using the land cover record of Landsat to track the change in impervious cover over time should provide a better understanding of its impacts on the weather and environment. It took a good deal of time to learn how to use the GS pan-sharpening, but it was rewarding to know that my small contribution to this project will help further the understanding of our environment when further research is conducted.

References:

Protocols for establishing a site and measuring tree cover based on the Globe Program protocols for Land Cover Sample Site (http://www.globe.gov/tctg/land_prot_samplesite.pdf?sectionId=209&rg=n&lang=EN) and Biometry (http://www.globe.gov/tctg/land_prot_biometry.pdf?sectionId=210&rg=n&lang=EN)

Landsat Data from USGS Global Visualization Viewer (http://glovis.usgs.gov/), Landsat Archive, Landsat 4-5TM

IKONOS Data provided by Global Land Cover Facility (http://glcf.umiacs.umd.edu/index.shtml)

Lillesand, T., & Kiefer, R. (1987). *Remote sensing and image interpretation*.New York: Wiley .

Hansen, M., DeFries, R., Townshend, J., Marufu, L., & Sohlberg, R. (2002). Development of a modis tree cover validation data set for western province, zambia. *Remote Sensing of Environment*, *83*, 320-335.

Shah, V., Younan, N., & King, R. (2008). Joint spectral and spatial quality evaluation of pan-sharpening methods. *Journal of Applied Remote Sensing*, *2*, 1-15