Etowah River Project Rapid Classification Tool

Natalee Dula

GISC 3200K – Programming for Geospatial Science & Technology

Dr. Huidae Cho

University of North Georgia

Abstract

The Etowah River is a popular recreational river that runs through Dawson Forest, a wildlife management area located in North Georgia. The location has been heavily damaged by nuclear activity and agriculture in the past and is now an area of interest for ecologists studying the long-term effects of the pollution and erosion caused by humans in the area. A project on freshwater invertebrates done in the 1950s is now being revisited to see what has changed in the time since the first collections were done, and it has been noted that the original sample locations are now experiencing heavy siltation. To see exactly how the area has changed since the original collections, comparing aerial and satellite imagery must be done. This can be a very time-consuming task as it requires processing a large amount of imagery; this toolbox aims to streamline and simplify the process for future students working on this project or similar projects.

Keywords: invertebrates, Etowah, siltation, change detection, satellite imagery

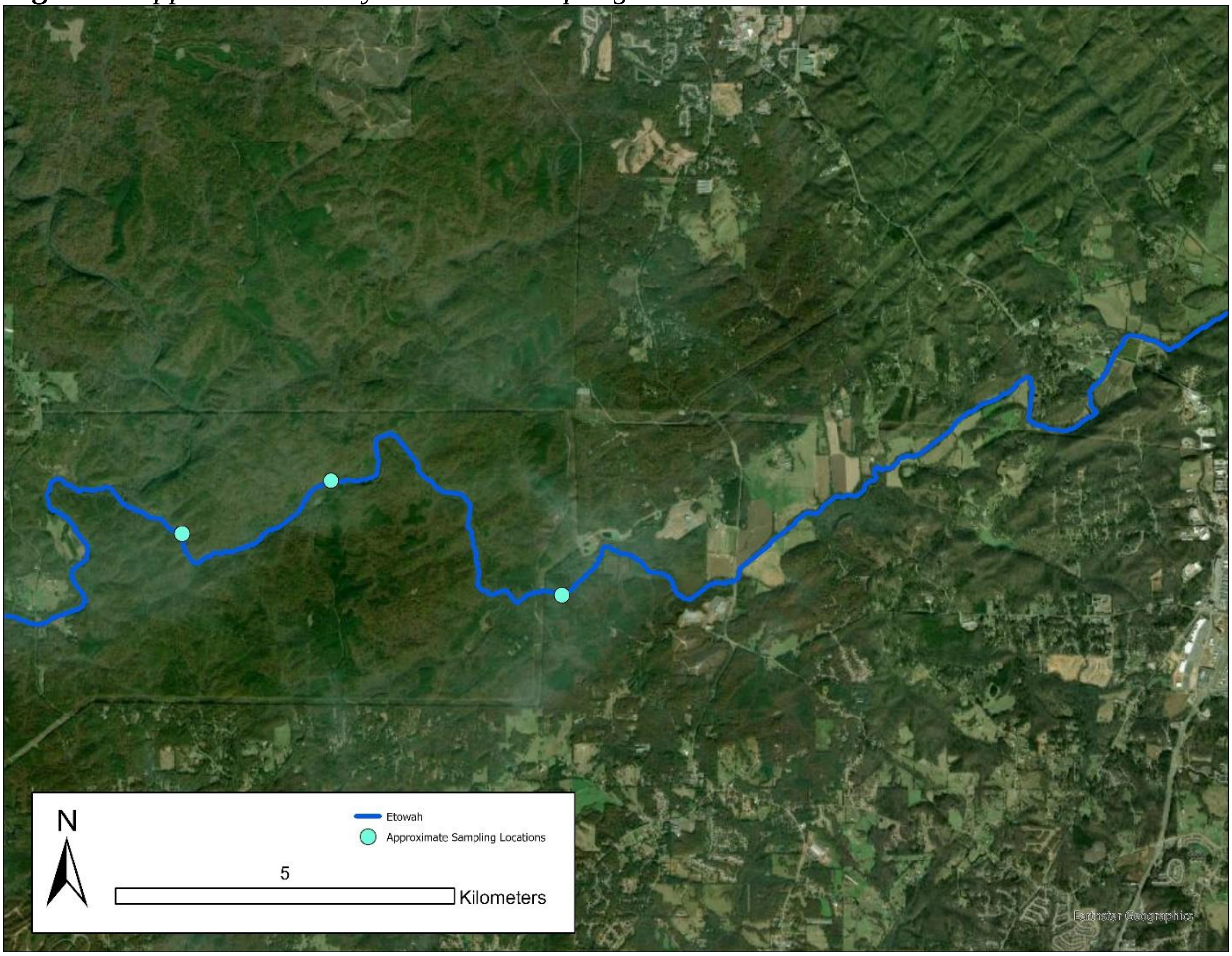
1. Introduction

The Etowah runs through several counties in Georgia, including Dawson county. Its size and biodiversity make it an important part of Georgia's ecosystem; however, it has had a lot of damage occur due to urbanization and other human involvement which has threatened many of the species living in the system (Wenger et al., 2010). In the 1950s, Dr. William Teitjen took on a project studying the invertebrates found at three locations in the portion of the river that runs through Dawson Forest. Funding was pulled from Dr. Teitjen's project, and it went without further work until Dr. Flood and a handful of UNG students took it on upon Dr. Teitjen's request to see how the invertebrates and stream have changed over the years after his work ended. During the 1950s, a nuclear facility was built in Dawson Forest, and prior to that, agricultural damage was already taking place; this is important to note, as the system was already being damaged well in advance of Dr. Teitjen's work and continued past it as well (Miller et al., 2019).

Despite the funding being pulled from the initial project, a "snapshot" collection such as Dr. Teitjen's can still prove helpful in future projects such as the current one, where the approximately 60 years of change can be viewed in the differences between the collections (Miller et al., 2019). One 2009 study of the Etowah heavily investigated land cover changes after the agricultural age of the area ended in the 1930s, and found that urbanization was a major negative indicator of macroinvertebrate diversity; overall, land cover changes accounted for up to 66% of macroinvertebrate variation (Walters, Roy and Leigh, 2009). This coincides with the work done by UNG's students; the invertebrates found during the most recent collections are consistent with those found in impaired aquatic systems and are overall less tolerant organisms than those that were found in the 1950s collections (Miller et al., 2019).

Ultimately, the goal of this project is to make a way to quickly create land use and change detection maps of the sampling locations and further upstream. Ideally, this will make it possible to go back many years for solid comparisons, but for the current project only 2019 and 2013 Landsat 8 data will be used for testing. Having many maps of land cover changes for this area across many years can potentially give insight into the gradual disappearance of less tolerant freshwater invertebrates in the area (Miller et al., 2019). Using an ArcPy toolbox can allow for more rapid processing of raster files and keep the tools needed for a project together and accessible for the user; the aim of this project is to create a toolbox that is easy for someone with minimal GIS experience to use so that newer students can take the project over without needing to take upper-level GIS courses first. This project is important as knowing how biodiversity of aquatic invertebrates is affected by nearby land cover changes can give insight into how that landcover change is affecting the water quality; this in turn can affect the economy, species diversity of an area, and human health, especially when involving a popular recreational river like the Etowah (Haque and Basak, 2017). Terrestrial biodiversity is also a concern, as damage to a waterway can heavily influence nearby non-aquatic animals, invertebrates, and plants that either rely on the water source or rely on food sources in the water; heavy terrestrial biodiversity changes can cause major impacts to the flow of local ecosystems (Newbold et al., 2015; Yuan, 2008).

Figure 1: *Approximate Study Area and Sampling Locations*



2. Materials and Methods

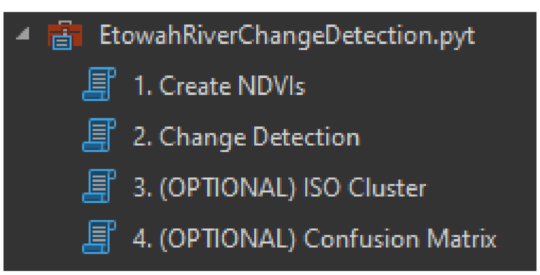
This project relies heavily on the availability of usable imagery of the area. While reviewing available imagery on the U.S. Geological Survey's EarthExplorer site, many of the older images were heavily covered in clouds which will make them virtually useless for land cover evaluation. As such, imagery in the GeoTiff format from the years 2013 and 2019 were selected due to their lack of cloud coverage and being within Landsat 8's range for consistency purposes; although the toolbox is built to be compatible with other Landsat imagery as well. In addition to imagery, the sampling locations do not have exact coordinates, and are instead approximated based on descriptions from Dr. Teitjen's original works; they are believed to be within 100 yards of the initial location. For land coverage alone this is not a very influential factor, however, it could marginally alter the types of organisms collected which may impact the perceived change in biodiversity.

The toolbox was meant to create an automated workflow in ArcPy to quickly prepare and process imagery for the project. The Composite Bands tool needs to be used before entering imagery into the tools, though future plans include incorporating this tool into the toolbox to automatically process rasters in a given folder. The toolbox currently has four tools incorporated;

the first is an NDVI tool that is functionally the same as the NDVI raster function with "Scientific Output" ticked but designed to process two rasters at once. The second tool creates the change detection imagery, achieved by using the Minus tool and subtracting the older NDVI from the newer NDVI. The third and fourth tools are optional based on what the users intents are; the third runs the Unsupervised ISO Cluster Classification tool and creates accuracy assessment points, while the fourth creates a confusion matrix using said accuracy assessment points after they have been manually corrected against the imagery or by visiting the coordinates and ground-truthing them. The fourth tool cannot be run without the third tool, as the confusion matrix requires the accuracy assessment points as an input. The confusion matrix itself is not necessary but allows the user to check the accuracy of the unsupervised classification against the raster imagery if desired. For a project such as this, it felt like a necessary addition. Similar projects have been done on a wider scale, such as one project over South California where creating land cover maps to detect forest change was automated in a similar manner in 2017 using varying Landsat data and an AutoLCD tool (Huang et al., 2017). Another project over Bangladesh coastal erosion examined land cover via random sampling over a series of years to aid in prediction of future land cover changes (Islam et al., 2020). Projects like these are helpful in building this project by reviewing what tools and methods they used and applying that knowledge to the more specified use of this project.

The decision to use a Normalized Difference Vegetation Index (NDVI) to view land cover changes seemed much more viable in this instance than just relying on classification methods, as other similar projects have used NDVIs in this manner (Aburas et al., 2015; Lunetta et al., 2003). There are many methods that can be used for a project such as this, but time and skill are factors that might inhibit the scope of it as many of the other projects had teams of experienced people working on them. Higher resolution imagery such as NAIP data may prove to be beneficial if available versus the current Landsat-8 imagery to ensure the highest level of accuracy, as many small-scale ecological land cover maps are difficult to properly work with at the 30 meter resolution (Nagel et al. 2014). This may not be feasible for the entire study area but is definitely something for future students to consider when collecting data to work with.

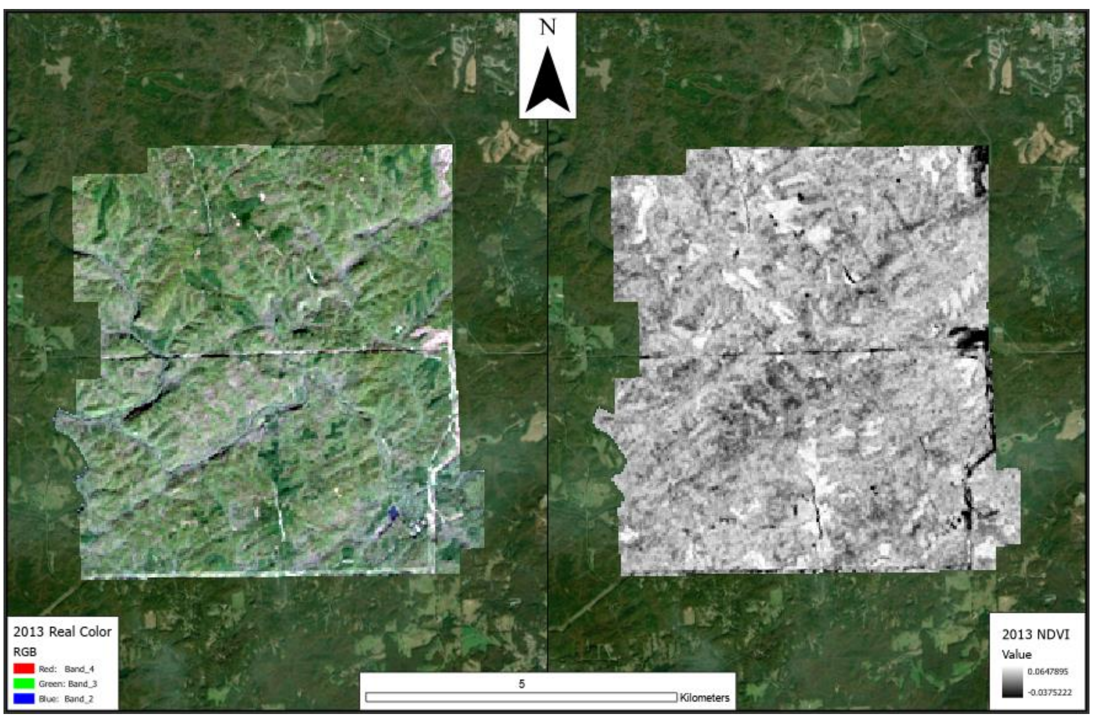
Figure 2: *Image of the toolbox*



3. Results & Discussion

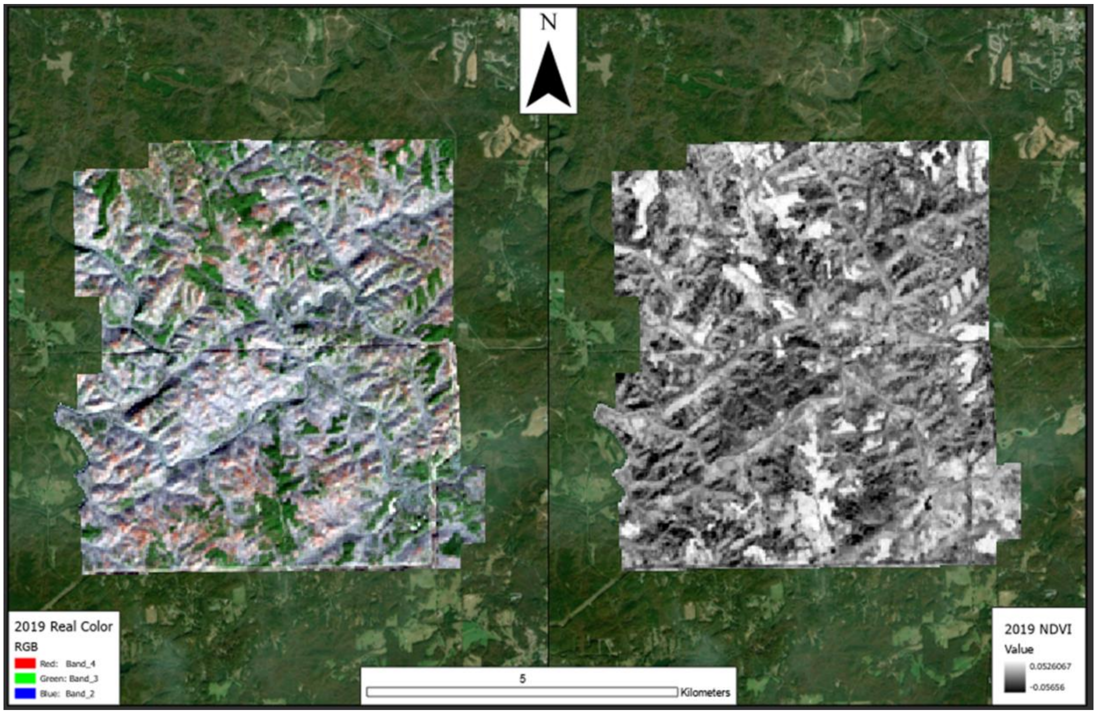
The first step in the process of using this toolbox involves using the Composite Bands tool on the downloaded Landsat-8 imagery to combine the bands into one raster. The band combinations of the imagery can then changed to match the Natural Color combination, which is 4, 3, 2 in Landsat-8, to show what the area looks like in real color. Other band combinations can also be used to view things like vegetation health, land vs. water, and urban development. The first tool can then be run to create an NDVI of two different years; the outputs will be in black and white but can be symbolized differently to portray the data better. The values for an NDVI tell what type of land cover is in the area, such as the type of vegetation, vegetation health, or substrate (Brown, 2013). Low or negative values can indicate barren or stony areas lacking vegetation, while moderate to high positive values indicate varying types of vegetation from sparse shrubbery to dense tree cover (Brown, 2013). These images can be useful in seeing how vegetation has increased or decreased in an area over time, and by subtracting one NDVI from another, a change detection image can be produced to see exactly what areas have gained (positive values) or lost (negative values) vegetation. NDVIs are primarily intended to look at vegetation changes, but in showing changes to vegetation, they also show other types of land cover changes. This is what makes them so beneficial for this sort of situation; they are easy for users to create and interpret and have very little room for human error. The unsupervised ISO cluster classification also has little room for human error but relies on high resolution and clear imagery to work very well. Supervised classification on the other hand needs more experience to use, is very time consuming to create training samples, and has a lot of room for human error—this being said, if there is enough time and someone with enough experience to do a supervised classification, then it would likely be better for something like this. Part of the intention was for something new user friendly however, so a supervised classification would simply not work in this instance.

Figure 3: *2013 Natural Color Composite Imagery vs NDVI*



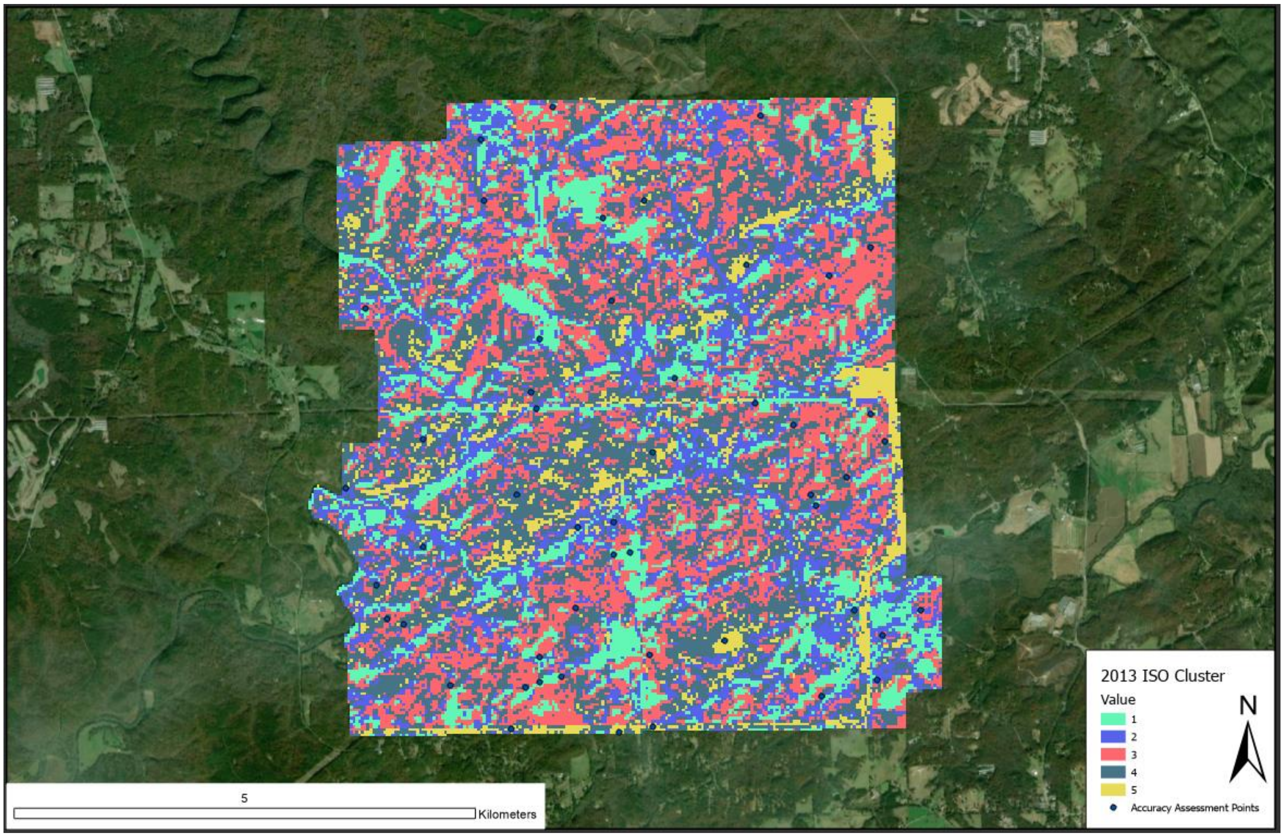
In this figure, you can see a lot of green vegetation; much of it appears to be low-lying vegetation such as shrubbery or grasses with few clumps of trees scattered throughout the area. The NDVI makes it easier to tell apart the stone and barren areas of soil from the vegetation, but on its own does not give much insight into the area. There are two distinct barren areas seen in both the real color map and the NDVI.

Figure 4: *2019 Natural Color Composite Imagery vs. NDVI*



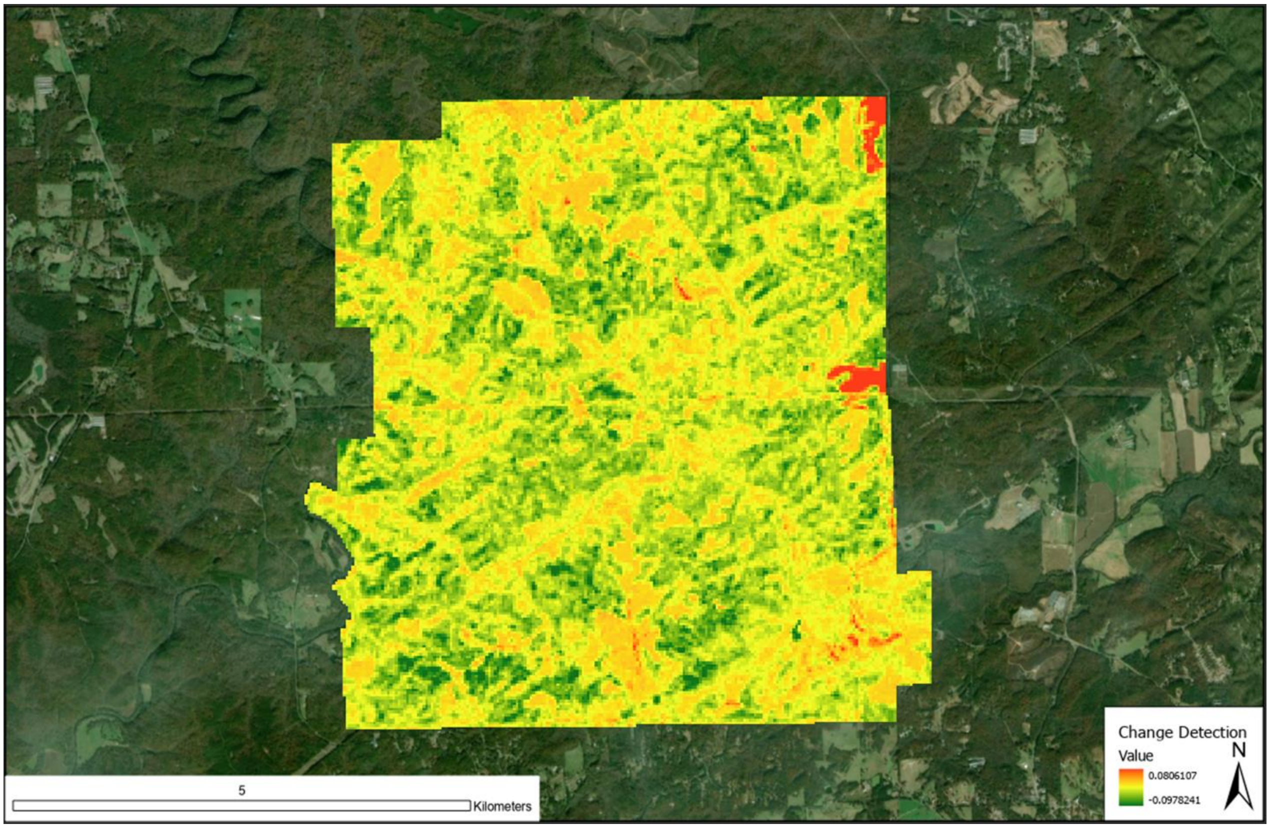
Comparing the 2013 imagery to the 2019 imagery, there appears to be far less low-lying vegetation with more dense clumps of trees visible amongst the barren areas. The barren and stony areas are much larger, indicating the imagery was probably taken in the fall or winter when the deciduous vegetation had died back, likely what can be seen in the green patches are evergreen trees. Ideally, comparing imagery from similar times of year gives the most accurate turnout when making a change detection map, because the vegetation growing at those times should be similar and not affected by deciduous versus evergreen coverage. For the sake of comparison and testing, two very different images were chosen to better exhibit the results of the change detection and its values. When using this toolbox for actual results, one will want to take care in getting clear imagery from a time of day with few shadows, no or little cloud coverage, and from a similar time of year as the image being compared to. Shadows and clouds both play a large role in if usable imagery is available or not, because both can majorly skew data.

Figure 5: *Example ISO Cluster Classification and Accuracy Assessment Points*



When the ISO cluster classification is performed, different numbers of classes can be defined. The number of classes created may be fewer than the defined number but may not exceed it; the capability to create more classes depends on how many similar pixels the program finds in the input feature. The symbology of the classified output feature will vary and usually not resemble the actual landcover classified and therefore will need to be manually symbolized to something more fitting and have the label names changed to match the land cover it is showing. Performing an accuracy assessment and confusion matrix is not necessary, but will ensure that the classification is accurate—the accuracy assessment points will individually need to be matched to a class against whatever the program believes is there, and then these updated points will be used to create a confusion matrix table. Ideally, the points will be corrected by ground-truthing via students visiting the actual location to see what land cover is there, however with old imagery that cannot be done and the points will instead need to be checked against the imagery used. The table will compare user accuracy to producer accuracy and give an idea of just how reliable the data is. An unsupervised classification only gives so much wiggle room with modifications to the classifications but can still be greatly helpful when looking at land cover types. With something like this that relies on accurate data, an accuracy assessment and confusion matrix would probably be desired if an ISO cluster classification is used in any part of the project.

Figure 6: *Change Detection (Symbolized)*



The change detection imagery is likely the most useful for the project, and the primary intended final product for the toolbox. Positive values indicate gain, while negative values indicate loss. The symbology has been changed to a red to green color-scheme to better visualize the data, which is something that can hopefully be incorporated into the toolbox to be done automatically in the future instead of the output image being black and white. The areas of loss are of course where the barren spaces were for the 2019 imagery, but interestingly there are two clumps of dense vegetation that are seemingly gained from 2013 to 2019 on the right side of the imagery. Reviewing more imagery in between the two years from similar dates could potentially give insight into these two spots of growth, as it seems strange that in 2013 the vegetation is not there but in 2019 it seems very lush and dense despite there being overall less vegetation in the 2019 map.

4. Conclusion

In conclusion, the toolbox could definitely have more useful tools and features added to it in the future with more time and experience. For the sake of Dr. Flood's project, this toolbox will likely be a great starting point for newer students looking to start in research. Future beneficial updates would include a batch Composite Bands tool, automated symbology for the NDVIs versus leaving them as black and white outputs, and perhaps the capability of running batch LiDAR analysis tools. The process for creating NDVIs and change detection maps has been streamlined and simplified as the original goal of the project entailed, but it could be pushed farther to provide more capabilities for this project or for similar projects that rely on the same tools.

Care has been taken to ensure the toolbox is flexible for more than just this project's data, and other Landsat imagery should work with it with no problems. Trying other types of imagery with the toolbox and ensuring compatibility would be an additional step for future work, as mentioned previously, NAIP or other high resolution rasters would up the accuracy of the land cover classifications and changes greatly. Any raster with both a red and near infrared band should likely work with the tool, as that is what the NDVI tool relies on to run, but this has not been tested as of yet.

The NDVIs are most likely going to be the best and most accurate data for what the intention of the toolbox is, but ISO cluster classifications can also be helpful when mapping land cover changes as it shows exactly what land cover is where (assuming the classes have been verified for accuracy) versus with the NDVIs, where it primarily just shows where vegetation is and is not. When paired with a runoff or erosion calculator, this toolbox should theoretically be able to answer nearly every relevant question when it comes to land cover changes pertaining to the Dawson Forest area and Etowah River where the original sampling took place.

5. References

Aburas, M. M., Abdullah, S. H., Ramli, M. F., & Ash'aari, Z. H. (2015). Measuring Land Cover Change in Seremban, Malaysia Using NDVI Index. *Procedia Environmental Sciences*, *30*, 238–243. https://doi.org/10.1016/j.proenv.2015.10.043

ArcPy, Esri.

Flood, Margi. Personal communication. (2020).

Haque, M. I., & Basak, R. (2017). Land cover change detection using GIS and remote sensing techniques: A spatio-temporal study on Tanguar Haor, Sunamganj, Bangladesh. *Egyptian Journal of Remote Sensing and Space Science*, *20*(2), 251–263. https://doi.org/10.1016/j.ejrs.2016.12.003

Huang, S., Ramirez, C., Kennedy, K., Mallory, J., Wang, J., & Chu, C. (2017). Updating land cover automatically based on change detection using satellite images: case study of national forests in Southern California. *GIScience and Remote Sensing*, *54*(4), 495–514. https://doi.org/10.1080/15481603.2017.1286727

Islam, M. M., Rahman, M. S., Kabir, M. A., Islam, M. N., & Chowdhury, R. M. (2020). Predictive assessment on landscape and coastal erosion of Bangladesh using geospatial techniques. *Remote Sensing Applications: Society and Environment*, *17*, 100277.

Landsat-8 imagery courtesy of the U.S. Geological Survey.

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., & Worthy, L. D. (2006). Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, *105*(2), 142–154. https://doi.org/10.1016/j.rse.2006.06.018

Miller, J., Pierce, G., Pesaresi, R., Tempia, R., Teitjen, W., & Flood, M. (2019). A COMPARISON OF BENTHIC INVERTEBRATES IN THE ETOWAH RIVER 1958 & 2018.

Nagel, P., J. Cook, B., & Yuan, F. (2014). High Spatial-Resolution Land Cover Classification and Wetland Mapping over Large Areas Using Integrated Geospatial Technologies. *International Journal of Remote Sensing Applications*, *4*(2), 71. https://doi.org/10.14355/ijrsa.2014.0402.01

Brown, J. (2013). NDVI, the Foundation for Remote Sensing Phenology, *US Geological Survey*. Available at: http://phenology.cr.usgs.gov/ndvi_foundation.php.

Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., … Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, *520*(7545), 45–50. https://doi.org/10.1038/nature14324

Walters, D. M., Roy, A. H., & Leigh, D. S. (2009). Environmental indicators of macroinvertebrate and fish assemblage integrity in urbanizing watersheds. *Ecological Indicators*, *9*(6), 1222–1233. https://doi.org/https://doi.org/10.1016/j.ecolind.2009.02.011

Wenger, S. J., Freeman, M. C., Fowler, L. A., Freeman, B. J., & Peterson, J. T. (2010). Conservation planning for imperiled aquatic species in an urbanizing environment. *Landscape and Urban Planning*, *97*(1), 11–21. https://doi.org/https://doi.org/10.1016/j.landurbplan.2010.04.006

Yuan, F. (2008). Land-cover change and environmental impact analysis in the Greater Mankato area of Minnesota using remote sensing and GIS modelling. *International Journal of Remote Sensing*, *29*(4), 1169–1184. https://doi.org/10.1080/01431160701294703