**NASA DEVELOP National Program**



Wise County Clerk of Court’s Office and NASA Langley Research Center

*Fall 2015*

Virginia Water Resources II

Utilizing NASA Earth Observations to Monitor the Extent of Harmful Algal Blooms in Lower Chesapeake Bay Watersheds

 **Technical Report**

Final Draft – November 17th, 2015

DEVELOP Wise County Clerk of Court’s Office

Arika Egan (Project Lead)

Jakub Blach

Zachary Tate

DEVELOP NASA Langley Research Center

Jessica Jozwik

Tyler Rhodes

Dr. Kenton Ross, NASA DEVELOP National Program (Science Advisor)

Bob VanGundy, University of Virginia’s College at Wise (Science Advisor)

Dr. DeWayne Cecil, Global Science and Technology, Inc. (Science Advisor)

Previous Contributors:

Sara Lubkin

Cassandra Morgan

# I. Abstract

Harmful Algal Blooms (HABs) in the Chesapeake Bay Watershed have an increasingly negative effect on the ecosystems in which they grow. They deprive their ecosystem of oxygen, produce harmful toxins, and mechanically damage other organisms. This disrupts the natural water chemistry and causes large-scale fish mortality events. Scientists from the Virginia Institute of Marine Science (VIMS) and Old Dominion University (ODU) monitor HABs and their effect on the water quality; however, the Chesapeake and its estuaries are geographically too large for the groups to continuously monitor HABs. This limits the group’s ability to monitor up-to-date locations of HABs and the water quality associated with them. To remedy this, surface reflectance data from Landsat 8 obtained from the USGS Earth Explorer, bathymetry imagery collected from NOAA CoastWatch, and *in-situ* data from VIMS and ODU were used to create a tool that produces a map of algal hotspots in the Chesapeake Bay area. Data were collected from August 17th, 2015. This tool will allow scientists at VIMS and ODU to identify the location of algal hotspots using current Landsat 8 data, as well as give them the ability to assess the timing, magnitude, duration, and frequency of HABs in the Chesapeake Bay Watershed.

**Keywords**

Remote Sensing, James River, York River, Landsat 8 OLI, Chlorophyll-a

# II. Introduction

The population of the Chesapeake Bay Watershed has doubled since 1950 and is currently approaching 18 million people (Chesapeake Bay Program 2015). As a result, increases in urban and agricultural land use have led to higher concentrations of nutrient runoff into the Chesapeake Bay and its estuaries (Ondrusek et al 2012). High concentrations of nitrogen and phosphorus in the water trigger the excessive growth of algae, known as Harmful Algal Blooms, or HABs (Ondrusek et al 2012). HABs have costly, negative impacts on water quality in the Chesapeake Bay. They form a film on top of the water, thus blocking sunlight necessary for photosynthetic activity in other beneficial organisms (Gilbert et al 2005). Additionally, as the algae die and decompose, they drastically decrease dissolved oxygen in the water (Gilbert et al 2005). By blocking sunlight and decreasing dissolved oxygen, entire ecosystems are negatively impacted and thrown out of balance by HABs.

HABs pose a threat to human health and well-being as well. They produce harmful, unpleasant smelling toxins that mutate underwater organisms and contaminate shellfish and oysters, making them unsafe for human consumption (Backer and McGillicuddy 2006). HABs often require beach closures to ensure safety. Communities also report massive fish mortality events during the bloom season from July to September, and oyster larvae are less likely to survive when HABs are present (Reese). As fishing, oyster harvesting, and tourism are three major economic industries for Virginia’s coastal communities, HABs pose serious threats to the economic security of the area (Chesapeake Bay Foundation 2015).

Actors on the local, state, and federal level are all concerned with the degraded water quality of the Chesapeake Bay. The Clean Air and Water Act of 1972 and President Obama’s Chesapeake Bay Executive Order of 2009 set water quality parameters and highlight the importance of restoring the health of the Chesapeake Bay. On the State level, a HAB Task Force comprised of representatives from the Virginia Department of Health, Virginia Institute of Marine Science (VIMS), Virginia Department of Environmental Quality (DEQ), the Marine Resource Commission, and Old Dominion University is tasked with identifying, monitoring, and researching HABs in an attempt to improve the water quality of the Chesapeake Bay. They focus on Virginia’s Chesapeake Bay, the James River, the York River, the Elizabeth River, and Mobjack Bay (Department of Environmental Quality 2015).

The HAB Task Force maintains 20 fixed testing stations throughout the region where various water quality parameters (chlorophyll-*a* content, salinity, temperature and turbidity), genetic molecular analysis, and HAB/phytoplankton identification tests are conducted monthly from May to November (Department of Environmental Quality 2015). Additionally, community members are encouraged to utilize a 24-hour HAB hotline that has been established for to report suspicious colors, smells, or fish kills (Virginia Institute of Marine Science 2015). When a HAB is detected or reported, the response team collects samples that are analyzed at different institutions depending on the nature of the report. Then, the VA Health Department determines future actions based on guidelines set by the Clean Water Act and State of VA Water Quality Standards (Department of Environmental Quality 2015).

 While these resources exist, the total area of the Bay is too large to continuously monitor and current methods do not allow for the desired real-time monitoring of the area. A more efficient and cost-effective method of identifying and studying HABs is necessary to assist local, state, and federal agencies and research institutions in their efforts to protect the Chesapeake Bay.

The objective for this project is to provide a method of identifying HABs in real time to the HAB Task Force using remote sensing technology. Building upon the work done by the Virginia Water Resources project from summer 2015 of NASA DEVELOP, we attempted to utilize Landsat 8 OLI images, and historical *in situ* data to create a python tool that highlights high concentrations of chlorophyll in the Chesapeake Bay and its estuaries. Because of available in situ data, we needed to modify our methods. Ultimately, our tool takes Landsat 8 OLI images, masks for land, cloud, shallow water, and turbid water, and produces two images highlighting chlorophyll based on the NDVI and combining bands 5, 4, and 3. Historically, researchers have utilized the Normalized Difference Vegetation Index to locate chlorophyll concentrations based on the amount of red and near infra-red light plants reflect (Shen et al 2012). Other band combinations, such as mid-infrared, near infrared, and red, have also been traditionally used to visualize chlorophyll (Horning 2004). Since vegetation absorbs nearly all red light, a red band can be helpful in visualizing chlorophyll. Near infrared helps differentiate between land and water. Since water absorbs nearly all light, the dark water is contrasted with bright reflectance of soil and vegetation on land. Finally, mid infrared is sensitive to moisture and is historically used to monitor vegetation. By combining these three bands, past researchers have been able to visualize chlorophyll concentrations on both land and water (Horning 2004).

Our tool applies these remote sensing concepts specifically to areas in the Chesapeake Bay. Our partners will be able to locate and monitor the timing, magnitude, duration, and frequency of HABs quickly and efficiently and identify areas of the Bay that need additional testing. This project addresses NASA’s Earth Science Water Resources application area and aligns with the goals of President Obama’s Chesapeake Bay Executive Order to target resources, define tools, strengthen scientific support for decision making, and develop focused and coordinated research programs to improve water quality of the Chesapeake Bay

# III. Methodology

 **Data Acquisition:***Landsat 8 OLI*
Landsat 8 surface reflectance data products were obtained from the United States Geological Survey’s (USGS) EarthExplorer for dates ranging from August 17th, 2015. This date was selected in order to combine it available *in situ* data also from August 17th, 2015. This date was also selected because it was a particularly cloudless day. The study area was virtually free of clouds. Path 15, Row 34 was used as the search criteria. A cloud mask, called the “CFmask” is provided from the USGS EarthExplorer when surface reflectance products are ordered. These images are the primary source of data. The pixel values of the images were used with *in situ* data in an attempt to find a regression model for identifying chlorophyll in the Chesapeake Bay.

*Ancillary Data*
*In situ* water sampling data were provided by Dr. Kim Reese from the Virginia Institute of Marine Science. Samples were obtained from a data collection cruise on August 17th, 2015, occurring between 10:00am and 12:00pm. The data cruise collection provided detailed measurements of various water quality parameters, including turbidity, temperature (degrees Celsius), *in vitro* chlorophyll measurements (µg/L), pH, and salinity (parts per thousand).

*Bathymetric Data*
Bathymetric data for the Chesapeake Bay at 30 meter resolution was downloaded as a DEM from the National Oceanic and Atmospheric Administration’s (NOAA) estuarine bathymetry website (<http://estuarinebathymetry.noaa.gov/bathy_htmls/M130.html>).

**Data Processing:** *Landsat 8 OLI*Originally, the pixel values in the Landsat 8 OLI data are given in integer values. This means that the outputs of any mathematical operations applied to them will default to integer values. Depending on the operation performed, resolution of the pixel values can be lost. Thus, true reflectance composites were compiled by dividing the pixel values by 10,000 for bands 1-7 of the Landsat 8 OLI data:

 $\frac{Integer band value}{10,000}$= $Rescaled floating point band value$

This converts the integer values into floating point values, and the resolution is conserved.

While the ending script produced by this product utilizes Landsat 8 OLI bands 2 – 5, this project provided processed images for all Landsat 8 OLI bands in the event the end users decide to use the processed data for additional analysis.

Next, land pixels were removed from each of the rescaled bands. This was achieved through the use of a normalized difference vegetation index (NDVI). The NDVI is calculated with the following formula: $\frac{Band 5-Band4}{Band5+Band4}$ = NDVI. Pixels with a value less than 0 correspond to water and were reclassified to 1. All other pixels are reclassified to “NoData”. This was done using the conditional evaluation tool (watermask = Con(ndvi,1,"","VALUE < 0"). This image was saved as “watermask.tif”. The watermask was applied to all rescaled floating point Landsat 8 OLI bands using the “ExtractByMask” tool in ArcMap. This removed pixels corresponding to land and pixels corresponding to water were extracted.

The cloud mask, or cfmask, provided by the Landsat 8 surface reflectance product download was used to remove clouds and cloud shadows. Table 1 describes the pixel values and their interpretation, as provided by USGS Product Guide for the Provisional Landsat 8 Surface Reflectance Product.

Using the conditional evaluation tool in ArcMap, pixels valued 0 or 1 were reclassified to “1”. All other pixels were reclassified to “NoData”. This image was saved as a cloudmask. The cloud mask was applied to the water-only rescaled bands using the “ExtractbyMask” tool in ArcMap.

Cannizzarro and Carder (2005) report stark differences in chlorophyll estimation between optically shallow and optically deep water. They developed a technique that classifies data as optically shallow or optically deep and created two different chlorophyll estimation algorithms based on that.

For the purpose of this project, it was decided that pixels corresponding to a depth of 2 meters or less would be removed. This was done for the second data analysis attempted. Shallow pixels reflect more light and produced what were believed to be false positives in our final product. Thus, a bathymetry mask was produced to remove these pixels. The mask was created using the conditional evaluation tool in ArcMap. Pixels with a value of -2 or greater were reclassified to “NoData” and all other pixels were reclassified to “1”. This image was saved as a bathymetrymask. It was applied to the land and cloud removed rescaled bands using the “ExtractByMask” tool in ArcMap.

Pixels corresponding to high sediment concentrations were removed for the second part of our data analysis. Like shallow water, it was suspected that high sediment concentrations were producing false positives when identifying areas of high chlorophyll concentration. This is removal is supported by Tebbs et. al. (2013). Lacaux et. al. (2007) describes a method for differentiating clear water from murky water, or water filled with sediment. This method is an index called the normalized difference turbidity index and is calculated using the following formula: $\frac{Red -Green}{Red+Green}=NDTI$. For Landsat 8 OLI, this becomes $\frac{Band4- Band3}{Band4 + Band3}=NDTI$. Somvansh et. al. (2011) uses the mean and the standard deviation of the NDTI to classify regions of high sediment, displayed in Table 2.

The standard deviation and mean of the NDTI were calculated from the NDVI image. Values corresponding to the “High” sediment classification were reclassified as “NoData” using the conditional evaluation tool in ArcMap. All other values were reclassified to “1”. This image was saved as and used to remove pixels with a high sediment concentration.

This final image, with land, cloud, shallow, and high sediment concentration pixels was used in our data analysis.

*Bathymetry*
The NOAA bathymetry data was available as three DEM tiles. The tiles were joined as a mosaic in ArcMap and saved as a .tiff. A bathymetry mask was created to remove pixels corresponding to a depth of 2 meters of less from the Landsat 8 OLI data. The mask was created using the conditional evaluation tool in ArcMap. Pixels with a value of -2 or greater were reclassified to “NoData” and all other pixels were reclassified to “1”. This image was saved as a bathymetrymask. It was applied to the land and cloud removed rescaled bands using the “ExtractByMask” tool in ArcMap.

*Ancillary Data*The VIMS data file came in an excel spreadsheet. It was imported into ArcMap. The coordinate system was set to the GCS\_WGS\_1984 geographic coordinate system. The data was exported to a shapefile. The pixel values of the processed Landsat 8 OLI bands were extracted to the VIMS shapefile using the “Extract Multi Values to Points” tool in ArcMap. Bathymetric values from the NOAA bathymetry mosaic data were also extracted to the VIMS shapefile.

There were a few data points from the VIMS shapefile that had no corresponding Landsat 8 OLI values. The missing pixels were most likely removed when shallow and cloud pixels were removed. The corresponding VIMS data values were removed by selecting all the points with a missing data in the Landsat 8 OLI data columns in the shapefile attribute table. Additional data points removed were those near any bridges and boats in the Landsat 8 OLI images. Data points near those features reflected extra light and produced artificially high reflectance values.

This collection of data was meant to be analyzed in the statistical analysis program R. To make the data readable in R, the VIMS shapefile was exported to a .dBASE. This .dBASE was opened in Microsoft Excel and then saved as a .csv. This .csv was used in R during the data analysis.

**Data Analysis:**
*First Method*The Landsat 8 OLI processed data used for this method had only land, cloud, and shallow pixels removed. A separate .csv file was created with the extracted Landsat 8 OLI data and the VIMS data. The .csv file with the processed Landsat 8 OLI and VIMS data cruise data was imported into R. A correlation analysis was run on the data, attempting to find a relationship between chlorophyll measurements and the processed Landsat 8 OLI band pixel values. 171 different regression models were run on the data, including linear-linear, linear-log, and linear-exponential relationships. The success of the correlations was determined by the R2 value. Table 3 shows the best models and regression formulas found in R.

As seen in Table 3, the best R2 value was 0.202, which is not a strong enough correlation to provide a good predictive model. A second analysis of the data in ArcMap provided evidence that the VIMS data cruise data and the Landsat 8 OLI data did not match up as they should.

Figure 1 shows the path of the VIMS data cruise overlaid on a color composition of the Lower York River. The color composition is made from the near infrared, red, and green bands of the processed Landsat 8 OLI data. This is a color composition commonly used to identify chlorophyll. For the data to be consistent with each other, the colors of the VIMS data cruise must match the colors on the Landsat 8 OLI composition. Three points of stark difference are circled in purple. This pointed to inconsistencies in the data that would prevent a strong correlation.

These inconsistent values were removed, leaving 500 data points that seemed to be consistent in the two data sets, and another attempt was made to find a regression model.

The same 171 models from the first regression round were used in R, but no strong correlation was found. The best R2 value produced was 0.35, which was not considered to be a strong correlation.

Due to time constraints arising when this data inconsistency was realized, a second, more visually based methodology was developed to produce chlorophyll concentration choropleth maps.

*Second Method*
The Landsat 8 OLI processed images for this method included the removal of high concentrations of sediment pixels. Two final choropleth maps were created. The first is a color composition consisting of the near infrared, the red, and the green bands of the processed Landsat 8 OLI data, created with the “Composite Bands” tool in ArcMap. This is a band combination commonly used to identify chlorophyll. Pixels colored a deep red correspond to high chlorophyll concentration; pixels colored a lighter red can correspond to mature or unhealthy growth; blues correspond to water, with lighter blues being shallower water and deeper blues being deeper water.

The second map was an NDVI map, created using the NDVI using the NDVI formula given previously. The values of the pixels are all negative, as expected when calculating an NDVI over water. However, values at the less negative end of the spectrum still refer to areas of high chlorophyll concentration.

# IV. Results & Discussion

*First Method*
This study attempted to find a regression model using *in situ* data and Landsat 8 OLI reflectance values. There are several reasons why this initial method did not work out. One possible explanation is the presence of currents and water vehicles in the bay at the time the data was taken. The Chesapeake Bay is host to a number of commercial and private boats, several of which can be seen in the Landsat 8 OLI images. The movement of these boats as they traveled throughout the Bay may have disturbed the location of the algal blooms between the times when VIMS collected the *in situ* data and when Landsat 8 flew over the Bay. Thus, the locations VIMS reported having high concentrations of algae may not have been where the Landsat 8 OLI data reported the locations. Figure 2 highlights a few boats captured in the study area in Landsat 8 OLI data from August 17th, 2015.

Another possible reason for the inconsistencies is the resolution of the Landsat 8 OLI data and that of the flow through data. Figure 3 shows a zoomed image of the 543 color composite with the VIMS data cruise displayed as dots. The size of each Landsat pixel is 30 m. The image shows a varying distribution of data points from the VIMS data within the Landsat 8 OLI pixels. The variability in the pixel distribution may have provided extra data points where none should have existed, thus skewing the true correlation of Landsat 8 OLI data and VIMS chlorophyll measurements.

Additionally, a discussion with representatives from VIMS let to the discovery that the equipment taking measurements of chlorophyll in the water is situated about a foot beneath the surface. Landsat 8 OLI data takes reflectance values right from the surface of the scene being imaged. An inconsistency then exists between the regions of the Chesapeake Bay that data is being collected from.

Furthermore, it’s possible that the models we attempted to find did not account for other factors which affect algae growth, like depth, salinity, and turbidity. While shallow water pixels were accounted for with the removal of pixels corresponding to a depth of 2 meters or less, the range of depths in the Chesapeake is dynamic, with a maximum depth of 174 feet (CBF). The objective of this project was to create a tool that estimates chlorophyll concentrations using new Landsat 8 OLI data for the end users. Thus, it was decided to begin the focus of the data processing on only Landsat 8 OLI data and the VIMS data cruise data.

*Second Method*
The second method involves the creation of two chlorophyll concentration maps: The 543 color composition map and the NDVI map. The images in Figure 4 display a side by side comparison of the two maps, which show the Lower York River. Each map displays similar regions of high chlorophyll concentration. Thus, the maps have similar abilities to provide the locations of regions of high chlorophyll concentration and harmful algal blooms.

Some differences in the visual displays in the images are important to point out. First, areas in light red on the 543 composition map are marked as bright red on the NDVI map. The large streak of dark red in the 543 color composition is also not entirely present in the NDVI image.

The NDVI image displays only relative vegetation concentration. This means that if the entire region contains a very low algal bloom, regions with concentrations at the higher end of the spectrum will be displayed as red. To better use this map, it’s necessary to have some kind of calibrating data to produce a prediction of what the range of chlorophyll values is. The best stretch for the NDVI map is also uncertain. It’s possible to display the map with an extremely polarized gradient, but it’s also possible to stretch is with a more linear and gradual color ramp.

The 543 composition image provides a visual difference between actively blooming chlorophyll (And thus algae) and mature or dying chlorophyll, through the different shades of red. This is potentially more useful than the NDVI image, as it allows users to identify locations of actively growing algal blooms versus blooms that have already grown. This is an important distinction to make as it allows the users of these maps to track the active causes of the active bloom, rather than arriving at a mature bloom and conducting guess work.

Since these differences exist and the maps are not calibrated with *in situ* data, it’s uncertain which map is better for identifying harmful algal blooms through high chlorophyll concentrations. Both are provided to the user so any reasoning they choose to use one over the other is up to their discretion.

The tool created from this project takes the raw Landsat 8 OLI data, goes through the data processing procedure outlined in the Data Processing section, and outputs two maps. The tool is called the Chesapeake Bay Chlorophyll Hotspot Identifier, or the CBCHI.

*Future Work*
Among the reasons the first method did not produce good results is the quality of the data. It’s necessary to obtain data that is more consistent throughout the set in order to find a good regression model. Another possibility is the size of our data set. Compared to the entire area this project was meant to address, the data available covered a fraction of the whole region. Figure 5 addresses the regions of the entire study area versus the data collection area.

One method that might prove interesting to try addresses the resolution of the Landsat 8 OLI data and the VIMS data cruise data. The method involves averaging the chlorophyll values of multiple cruise data points in one pixel of the Landsat 8 OLI. This would be an attempt to reduce the potential skewing of the regression model due to the inconsistent distribution of the original Landsat 8 OLI and cruise data. This would produce one cruise data point per Landsat 8 OLI pixel and might produce a better regression model.

# V. Conclusion

The objective of this project was to create a tool that takes in Landsat 8 OLI data and produces a map providing estimations of chlorophyll concentration for the Chesapeake Bay Watershed. This tool was to be created in python. It would first process the Landsat 8 OLI data by removing land pixels, cloud pixels, and pixels corresponding to a depth of 2 meters or shallower. The tool would be incorporated with a regression model obtained by the correlation of Landsat 8 OLI surface reflectance images and *in situ* data provided by the Virginia Institute of Marine Science from a data cruise along the Lower York River. The tool would apply the regression model to the processed Landsat 8 OLI data and produce a map with predictions of chlorophyll concentrations.

The values of the data were inconsistent between Landsat 8 OLI and the VIMS cruise data. Thus, a model producing chlorophyll estimations could not be produced. The python tool instead produces two maps meant for highlighting relative chlorophyll concentration. The first map is a 543 color composition of processed Landsat 8 OLI data. The second map is an NDVI map created from bands 5 and 4 of the processed Landsat 8 OLI data.

While the tool provides maps with visual representations of chlorophyll concentration instead of numerical predictions, it is a good first step into identifying HABs in the Chesapeake Bay. There is currently no reliable method of real-time monitoring of HABs in the entire Chesapeake Bay. This tool provides maps of the entire Chesapeake Bay, allowing users of the CBCHI to have access to information the size of the bay normally precludes.

**VI. Acknowledgments**The Virginia Water Resources II team would like to thank the following individuals for their assistance in the research and development of this project and the Chesapeake Bay Chlorophyll Hotspot Identifier:

Science Advisors:

* Dr. Kenton Ross, NASA DEVELOP
* Robert VanGundy, University of Virginia’s College at Wise
* Dr. DeWayne Cecil, Global Science and Technology, Inc.

Partners:

* Dr. Kim Reece, Virginia Institute of Marine Science
* Russ Baxter, Virginia Deputy Secretary of Natural Resources for the Chesapeake Bay
* Will Hunley, Hampton Roads Sanitation Department
* Todd Egerton, Old Dominion University

Others:

* Dr. Sarah Lubkin, University of Mary Washington (previous contributor)
* Cassandra Morgan NASA DEVELOP (previous contributor)
* Melanie Salyer, NASA DEVELOP (Wise County)

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract NNL11AA00B and cooperative agreement NNX14AB60A.

# IV. Appendices

**Table 1**

|  |  |
| --- | --- |
| **Pixel Value** | **Interpretation** |
| 255 | Fill |
| 0 | Clear |
| 1 | Water |
| 2 | Shadow |
| 3 | Snow |
| 4 | Cloud |

Table 1. Description of the pixel values and their interpretation of the cfmask, as provided by the USGS Product Guide for the Provisional Landsat 8 Surface Reflectance Product.

**Table 2**

|  |  |
| --- | --- |
| **Sediment Classification** | **Formula** |
| Low | Mean – Standard Deviation |
| Moderate | Mean + Standard Deviation |
| High | More than Moderate |

Table 2. Regions of sediment concentration, as defined by Somvanshi et. al (2011).

**Table 3**

|  |  |  |  |
| --- | --- | --- | --- |
| Formula | R2 | formula | R2 |
| ~exp(b1/b2) +b1+ b2+ b3+b4 | 0.1868 | ~exp(b4/b3) +b1+ b2+ b3+b4 | 0.1889 |
| ~exp(b1/b2) +b1+ b2+ b3+b5 | 0.1796 | ~exp(b4/b3) +b1+ b2+ b3+b5 | 0.1884 |
| ~exp(b1/b2) +b1+ b2+ b3+b6 | 0.1613 | ~exp(b4/b3) +b1+ b2+ b3+b6 | 0.1905 |
| ~exp(b1/b2) +b1+ b2+ b3+b7 | 0.1621 | ~exp(b4/b3) +b1+ b2+ b3+b7 | 0.1902 |
| ~exp(b1/b2) +b1+ b2+b3+b4+b5 | 0.1872 | ~exp(b4/b3) +b1+ b2+b3+b4+b5 | 0.1893 |
| ~exp(b1/b2) +b1+ b2+b3+b4+b6 | 0.1878 | ~exp(b4/b3) +b1+ b2+b3+b4+b6 | 0.1909 |
| ~exp(b1/b2) +b1+ b2+b3+b4+b7 | 0.1877 | ~exp(b4/b3) +b1+ b2+b3+b4+b7 | 0.1907 |
| ~exp(b1/b2) +b1+ b2+b3+b4+b5+b6 | 0.1959 | ~exp(b4/b3) +b1+ b2+b3+b4+b5+b6 | 0.2011 |
| ~exp(b1/b2) +b1+ b2+b3+b4+b5+b7 | 0.1962 | ~exp(b4/b3) +b1+ b2+b3+b4+b5+b7 | 0.2009 |
| ~exp(b1/b2) +b1+ b2+ b3+b4+b5+b6+b7 | 0.1968 | ~exp(b4/b3) +b1+ b2+ b3+b4+b5+b6+b7 | 0.202 |

Table 3. Table of attempted regression models in the statistical program R. R2 value was 0.202, which is not a strong enough correlation to provide a good predictive model



**Figure 2.** A true color composition of the Lower York River, created from Landsat bands 4, 3, and 2. Circled in red are three boats that were traveling in the river at the time the Landsat 8 OLI captured the image. The movement of boats like these in the water may have shifted the location of algae in the water, and thus affected the locations of detected chlorophyll in the Landsat 8 OLI data and the VIMS data cruise data.

**Figure 1.** Image of a 543 color composition used for displaying chlorophyll levels overlaid with the VIMS data cruise path. The data points are colored according to chlorophyll level, with green being low and red being high. The four circled areas highlight inconsistencies in the image and the cruise data, including inconsistencies in measurements where the data overlaps, where chlorophyll exists in the image but not in the cruise data, and vice versa.

**Figure 4.** These images are the final results of the code created for this project. The image on the left is a 543 color composition meant to display chlorophyll. Dark red pixels correspond to areas of high chlorophyll and light red pixels correspond to mature or dying vegetation. Light blue pixels correspond to shallower water and dark blue pixels correspond to deeper water. The image on the right is an NDVI with a 20% min percent clip and a color ramp applied to it. Pixels in red correspond to areas of relative high concentrations of chlorophyll and areas in green correspond to areas of relative low concentrations of chlorophyll.

**Figure 3.** Close up of the VIMS data cruise data on a 543 color composite of fully processed Landsat 8 OLI data. The boxed pixel at the top has 3 data points in it, while the bottom boxed pixel has 18 data points in it. The amount of data points in each pixel is a result of the speed of the boat collecting the data. The excess and dearth of cruise data possibly skewed the data used for finding a good regression model.