**NASA DEVELOP National Program**



NASA Langley Research Center

*Fall 2015*

North Carolina Ecological Forecasting II

Update of NOAA C-CAP Wetland Delineation and Further Disaggregation of Land Use Classes using Remote Sensing

 **Technical Report**

Final Draft – October 8, 2015

Ben Roberts-Pierel (Project Lead)

Brett Buzzanga

Michelle Pasco

Jake Patrick

Benjamin Charlem

Taylor Sage

Emily Adams, NASA DEVELOP (Advisor)

Dr. Kenton Ross, NASA DEVELOP (Science Advisor)

Previous Contributors:

Zand Bakhtiari

Stephen Zimmerman

Kayla Patel

Brad Gregory

# I. Abstract

This project focused on ecological forecasting for wetlands in the Albemarle-Pamlico watershed in Northeastern North Carolina and Southeastern Virginia. The Albemarle-Pamlico watershed encompasses the second largest estuary system in the United States. Understanding land cover types and uses is incredibly important in managing the myriad of uses for, and stressors on, this valuable resource. In partnership with the Albemarle-Pamlico National Estuary Partnership (APNEP), this project aimed to provide an updated version of NOAA’s Coastal Change Analysis Program (C-CAP) land-use classification, with a specific focus on delineation of wetland types within this watershed. The project also further disaggregated land cover types such as the invasive species, *Phragmites australis*. The team utilized a supervised land classification methodology and cross-referenced Landsat 8 imagery with ground truth, LiDAR or Digital Elevation Models (DEM), the National Hydrological Dataset (NHD) and soil datasets to create inputs for a Random Forest Model in Google Earth Engine. The end goal of the project was to produce maps and a methodology by which APNEP can continually update wetland types and to establish if there is a correlation between wetland type and wetland health within the watershed. This was all with the aim of helping APNEP to make informed policy and management decisions.

**Keywords**

Albemarle, Pamlico, Landsat 8, Satellite, Supervised Classification, Spatial Analysis, Estuary

# II. Introduction

**Background Information:** The Albemarle-Pamlico region is the ‘largest lagoonal estuarine system in the United States’ (EPA, 2007). It consists of six river basins (Fig 1) as well as the Albemarle and Pamlico Sounds. Wetlands provide enormous value to the Southeastern United States through numerous ecosystem services. Wetlands absorb storm surges and strong winds; they minimize eutrophication by retaining and filtering agricultural and urban runoff; and they provide protection for developing fish and amphibian species, thereby continuing the existence and proliferation of terrestrial and oceanic biodiversity (APNEP, 2012).

However, the complex and dynamic nature of the Albemarle-Pamlico estuary network makes it a delicate system. Climate change induced sea level rise and land use change, secondary to population growth, are two significant threats to the Albemarle - Pamlico region (Carpenter, 2012). Sea level is currently rising at a faster rate than wetland vegetation can keep pace with, which leads to inundation and erosion at about -3.3 feet/year (Carpenter, 2012). In the twenty years between 1990 and 2010, the Albermarle - Pamlico region’s population grew by 36%, from 2,762,409 to 3,756,019 (APNEP, 2012). This population increase necessitates land change in the form of urban development and agricultural expansion. Even when wetlands are not specifically drained and converted to these purposes, mismanagement of natural resources, waste production, and nutrient runoff have reduced the water quality and human health in the Albemarle-Pamlico watershed (Carpenter, 2012).

In order to protect this region, the dynamics and characteristics of the estuary system must be better understood. A previous stage of this investigation, utilizing remote sensing data and indices that measured the extent and health of wetlands, concluded that wetland health has been deteriorating over time. However, to fully grasp the implications of this deterioration we must have clear delineations on the types of wetlands that currently exist, which types are deteriorating, and how this will affect the region as a whole.

**Project Objectives:** The Albemarle-Pamlico National Estuary Partnership (APNEP) currently relies on land use and land cover (LULC) data from NOAA’s Coastal Change Analysis Program (C-CAP) to understand land use designations and wetland types. This information helps to inform their management decisions by enabling them to decide where to allocate resources and to understand temporal trends. However, C-CAP is only updated every five years and is designed for regional coastline classifications. The team aimed to produce an updated classification that was tailored to the Albemarle-Pamlico estuary by providing additional land classification types. Within this classification, the project aimed to identify the invasive species *Phragmites Australis*, a focus of many management efforts. The objective was not only to provide APNEP with an updated set of images based on most recent available Landsat 8 imagery, but also to create a method of replicability so it would be possible for APNEP to update these classifications on a more regular basis than what C-CAP offers.

**Study Area**: The Albemarle-Pamlico region is located on the coast of Northeastern North Carolina and Southeastern Virginia. The basin consists of 31,000 square miles of land and water, including two major Sounds, Albermarle and Pamlico, and seven major river basins (Fig 1), six of which APNEP monitors. Three major land cover classes—forests (40.1%), croplands (25.3%), and wetlands (15.8%)—account for most of the LULC in the basin (APNEP, 2012; Carpenter, 2012).



Figure : Map of the Albemarle-Pamlico study area: river basins and major sounds.

**Study Period:** The project covered the summer of 2015.

**National Application Area:** This project focuses on and contributes to the NASA Applied Sciences Application Area of Ecological Forecasting through the delineation of current wetland extent using remote sensing data, and software such as ArcGIS an ERDAS IMAGINE. Although use of Landsat data for LULC is fairly common, this study provided the partner organization with a useful tool for the specific region of focus and ideally a methodology that will be useful not just to APNEP but potentially other areas of the Applied Sciences program at NASA or to the public. The integration of other data from USGS and NOAA, as well as comparisons with the results from the previous portion of the project, represent other processes, which could lay groundwork for future projects. This methodology could contribute to the Application Area by increasing uptake and utilization amongst groups outside NASA by showcasing datasets, their uses, and the accessibility of the data NASA produces.

**Project Partner:** Partnership with APNEP started in the spring of 2015. The primary goal of APNEP is to manage, conserve and protect the resources of the Albemarle-Pamlico estuary system. In order to do this, they require a better understanding of how the discrete components of the ecosystem function together as a whole and what stressors are adversely impacting this resource. They are focused on analyzing both short and long term trends and utilizing citizen’s groups, researchers and federal agencies, as well as governments from the local to federal level to accomplish their objectives (APNEP, 2012).

# III. Methodology

This study conducted a supervised land classification by using elevation, hydrography, soil, and crop data to select training sites for a Random Forest and a Maximum Likelihood Land Classification Model (Table 1). The Random Forest model runs in the open-source program Google Earth Engine (GEE) and the Maximum Likelihood classification runs in ESRI ArcGIS. Both methods convert spectral radiance values as recorded in Landsat 8 images to user - specified land cover types through two different statistical processes.

|  |  |  |
| --- | --- | --- |
| **Dataset** | **Source** | **Type/Specific Item** |
| Landsat 8 Operational Land Imager (OLI) | NASA EOS/USGS | Path 14, Row 35  Path 14, Row 36 |
| Landsat 8 Climate Data Record (CDR) | NASA EOS/USGS | Path 14, Row 35  Path 14, Row 36 |
| National Hydrography Dataset (NHD) | USGS | Waterbodies |
| Coastal Change and Analysis Program (C-CAP) | NOAA | 2010 iteration |
| Soil Survey and Geographic Database (SSURGO) | USDA |  |
| Digital Elevation Models (DEMs) | USGS | 1/3 arc-second and 1 arc second |
| LiDAR | NOAA | 2001 NC Phase I |
| Cropscape Annual Crop Data | USDA | 2010-2014 growing seasons for area constrained by the Landsat scenes |

Table 1: This provides a brief overview of the data utlized in the completion of this project. Federal agency titles have been abbreviated for the purposes of space. Please consult Appendix I for a list of acronyms.

Some of these data required preprocessing. Landsat 8 OLI scenes were preprocessed by the USGS to convert pixel values to surface reflectance as well as correct for atmospheric interference in the Provisional Landsat 8 Surface Reflectance data product, part of the Landsat 8 Climate Data Record (CDR) dataset (USGS, 2015). Images from path 14 row 35 and three images for path 14 row 36 were obtained for three dates, (June 23, 2015, July 25, 2015, and August 10, 2015) for a total of five Landsat scenes. The scenes were chosen due to the relatively high quality of the images: little cloud cover and most recent peak vegetation.

Processing and analysis proceeded by aggregation of these datasets in ESRI’s ArcGIS software. Eight 1/3 arc-second (10m resolution) DEMs were mosaicked to create one representative raster of the area covered by the Landsat scenes. The ‘swamp/marsh’ type of the NHD Waterbody feature class was clipped to the study area. The “Depth to Water Table” and “Hydric Rating” SSURGO attributes were similarly merged for analysis within ArcGIS. The former attribute described the depth to the water table, given in centimeters; the latter demarcated the percentage of an area’s soil that qualifies as hydric, a soil commonly found in wetland areas.

These datasets were compared in order to select polygons, the training sites, which served as calibration points for the classification models. Using ArcGIS and ESRI Imagine, training sites were chosen that clearly delineated a specific land cover type. The C-CAP classifications (Appendix II) guided the choices in land cover types, but some regionally inappropriate classes were ignored (such as tundra), while others of local concern such as crop types and invasive species were added (Appendix III).

In identifying *Phragmites Australis*, the process relied heavily on Google Earth and georeferenced photos within that program. These locations where *Phragmites* was positively identified were then compared with their respective points on a Landsat composite. Composites were tailored to best visualize differences in vegetation (bands 6 or 7, 5 and 4 or 2), and an approximate visual signature for clumps of *Phragmites* was identified.

The training sites encapsulated a minimum of 30-40 pixelsto allow for a margin-of-pixel-overlap in the Landsat images (30 m2 resolution). 30-50 training sites were selected for each land cover classification type. The sites and chosen land use classes were then input to Google Earth Engine, which extrapolated the classification of the training sites to the entire spectral area (2 Landsat scenes), generating an image of land classification for the whole study area.

**Accuracy Assessment**

The accuracy assessment was carried out using ERDAS Imagine and Microsoft Excel and points were verified using ancillary datasets in ArcGIS and imagery in Google Earth. First, the Accuracy Assessment tool in Imagine was used to generate approximately 30 random points per class. As a result of time constraints the points were generated based on a version of the C-CAP imagery which was re-classed to the 16 classes used in the final product. The C-CAP imagery emphasizes some classes very differently than our classification, Estuarine Scrub/Shrub Wetland for example, so additional points for this category were added to test this class. Each of the points were then analyzed manually and given a reference class using a combination of datasets utilized in the selection of training sites, as well as Google Earth imagery and georeferenced photos. Points that were unclear or outside the study area were discarded. This yielded 524 total points.

The Points to Values tool in ArcGIS was used to generate class pixel values at the assessment points. The known (accuracy assessment) points were then compared with the pixel values taken from the final classification in excel and a percent accuracy was calculated.

# IV. Results & Discussion

The principal result of the project was the LULC map produced from the training sites assembled in ArcGIS (Figure 2). Although classifications were produced in ArcGIS and GEE, the GEE result was a more representative result and therefore served as the basis for our analysis.



Figure Final classification produced in Google Earth Engine

As noted in the methods, an accuracy assessment was run on this image, which resulted in an accuracy of about 48 percent. In an ideal scenario, the reference points used would be ground truth points from randomly selected locations. However, without this as an option, as much care as possible was taken in attributing the points, but this method clearly resulted in a fairly high degree of residual error.

Although the accuracy was fairly low across the entire image, identification of certain classes was much more successful. A confusion matrix would be representative of this phenomenon but unfortunately due to time constraints and technical difficulties this is not included here. Therefore, it is more illustrative to look at specific regions or features within the overall classification than the finished product as a whole. This is seen in Figure 3 below where images of the Great Dismal Swamp in North Carolina exemplify stark differences in the classifications.

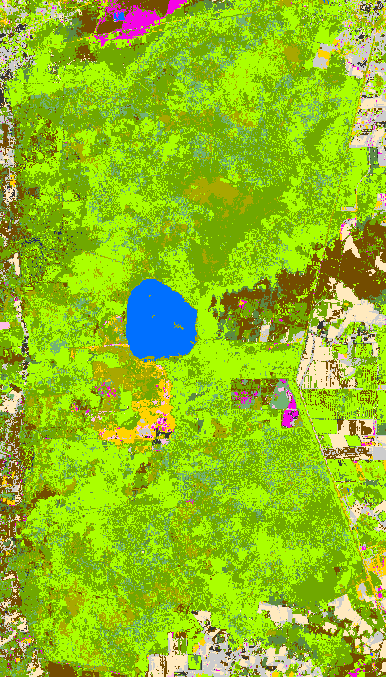
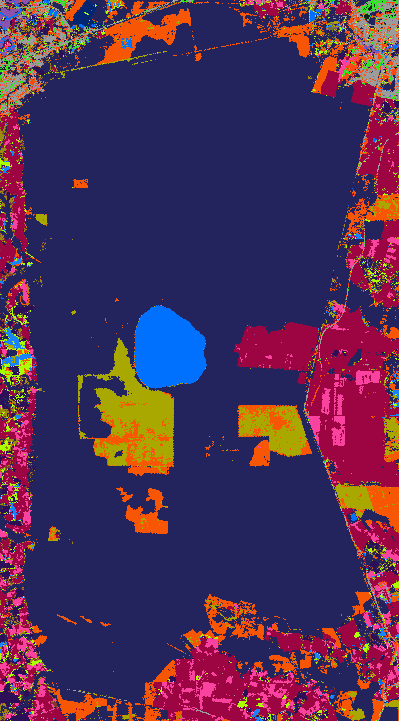


Figure Series of sample images of the Great Dismal Swamp from project classifications including the original CCAP (left), results from Google Earth Engine (Center) and results from ArcGIS (right). All images utilize the same legend color scheme.

One of the principle uses of these results was to contrast with the C-CAP product to highlight areas where our classifications differ, particularly in the wetland classes (Figure 4). The most prevalent differences found were in the classification of palustrine scrub/shrub wetlands by C-CAP, when in reality ground truth data show this area should be largely defined as palustrine forested wetlands. Scrub/Shrub wetland is defined by vegetation under 5m in height while forested wetland is characterized by vegetation which exceeds 5m. This result was verified with a trip to the Great Dismal Swamp where photos and GPS points were collected (Figure 5).

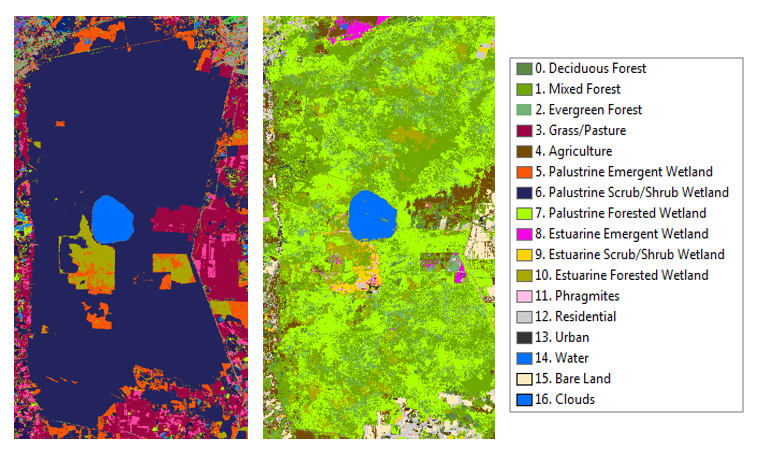


Figure left image is from CCAP and depicts an area in the Great Dismal Swamp where CCAP classifies Palustrine Scrub/Shrub Wetland when in reality it is Palustrine Forested Wetland as depicted in the final classification (right).



Figure 5 Photo of the forested wetland in the Great Dismal Swamp. Photo was taken at the GPS point labeled in a red circle on the images in Figure 4 above.

One of the other goals of this project, was to attempt to highlight areas where the invasive species *Phragmites Australis* might be found. Although certain verification would require ground truthing, the team was able to identify *Phragmites* in certain areas through a combination of spectral signatures in Landsat 8 band composites and Google Earth as well as georeferenced photos.

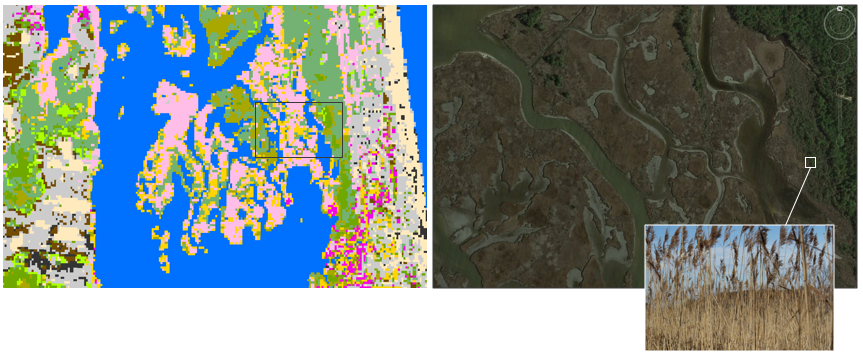


Figure 6 Image displays a snapshot of the final classification (left), with Phragmites Australis shown in pink. The other half is a sample from Google Earth (right) with a georeferenced photo of Phragmites Australis.

There are a few potential sources of error in this project. The first and likely most influential is human error. The team developed a methodology for identifying wetland class types to attempt to standardize the process, however, selection of training sites remains largely subjective to the analyst. Additionally, the process leveraged a wide range of data to cross reference, and any error in these layers could also present a compounding factor. There is also a high level of potential error in the classification mechanism when trying to identify classes which have such a similar spectral signature (e.g. forest and forested wetland). In a similar realm, influences of factors such as clouds have an impact on the classification and how the algorithm is able to identify classes on the ground.

# V. Conclusions

The team was able to complete a series of classifications and identify areas where C-CAP can be improved. This was accomplished through the use of a variety of NASA, USDA and USGS-derived datasets. However, it would have been greatly improved by ground truth points not available to us during the study. In order to make this portion of the assessment more statistically valid, a much wider ranging study would need to be conducted. The scope of this work was sufficient for highlighting specific areas of focus but the work involved in creating a representative model over the entirety of the study area would need to be much more in depth.

Based on the statistics obtained from the accuracy assessment, improvement in the selection of training sites or the methodology used could be improved. Going forward, it would be advisable to incorporate more ground truth data into an accuracy assessment or even the classification stage. This kind of data would be very useful in having a broader conceptual idea of what different types of wetlands look like and where one might expect to find them. Although we were able to collect a very small number of points, they were not enough to represent anything statistically significant.

It would also be helpful to expand the scope of this project in the data collected and the robustness of the methods used. This could involve the network of citizens, scientists, policymakers and industry which APNEP levies for some of their other work. There is no doubt that parts of the Albemarle-Pamlico basin are being threatened by stressors from land use changes and human activity and knowing more about where this is happening and which specific wetland types are most threatened should be of use in ameliorating these issues in the future.

# VI. Acknowledgments

Dr. Kenton Ross – National Science Advisor, NASA DEVELOP National Program

Emily Adams – Center Lead LaRC, NASA DEVELOP National Program

Rebekke Muench – Assistant Center Lead LaRC, NASA DEVELOP National Program

Bill Crowell – Ablemarle-Pamlico National Estuary Partnership

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.

# VII. References

Albemarle-Pamlico National Estuary Partnership, 2012. Comprehensive Conservation and Management Plan 2012-2022. *APNEP*. Available at: http://portal.ncdenr.org/c/document\_library/get\_file?uuid=e6600731-daed-4c5f-9136-253f23c9bbcf&groupId=61563 [Accessed October 7, 2015].

Carpenter, D., 2012. *Albemarle-Pamlico Ecosystem Assessment 2012*, Raleigh.

Environmental Protection Agency, 2007. National Estuary Program Coastal Condition Report. *EPA*. Available at: http://water.epa.gov/type/oceb/nep/upload/2007\_05\_09\_oceans\_nepccr\_pdf\_nepccr\_nepccr\_se\_partb.pdf [Accessed October 7, 2015].

Norning, N. & Wegman, M., 2013. Script for Classification of Satellite data [in R].

United States Department of Agriculture, Hydric Soils - Introduction. *USDA*. Available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/use/hydric/?cid=nrcs142p2\_053961 [Accessed October 7, 2015].

United States Geographical Survey, 2015. Product Guide: Provisional Landsat 8 Surface Reflectance Product. *USGS*. Available at: http://landsat.usgs.gov/documents/provisional\_l8sr\_product\_guide.pdf [Accessed October 8, 2015].

# VIII. Content Innovation

1. *Phragmites Australis* identification

# IV. Appendices

**Appendix I**

|  |  |
| --- | --- |
| **APNEP** | Albemarle-Pamlico National Estuary Partnership |
| **C-CAP** | Coastal Change Analysis Program |
| **DEM** | Digital Elevation Model |
| **LULC** | Land Use Land Cover |
| **NHD** | National Hydrography Dataset |
| **NOAA** | National Oceanic and Atmospheric Administration |
| **NPO** | National Project Office |
| **OLI** | Operational Land Imager |
| **SSURGO** | Soil Survey and Geographic Database |
| **SWIR** | Albemarle-Pamlico National Estuary Partnership Short-Wave Infrared |
| **TOA** | Operational Land Imager Top of Atmosphere |
| **USDA** | United States Department of Agriculture |
| **USGS** | United States Geological Survey |

**Appendix II**

C-Cap Attributes:

0 Background,

1 Unclassified (Cloud, Shadow, etc),

2 High Intensity Developed,

3 Medium Intensity Developed,

4 Low Intensity Developed,

5 Developed Open Space,

6 Cultivated Land,

7 Pasture/Hay,

8 Grassland,

9 Deciduous Forest,

10 Evergreen Forest,

11 Mixed Forest,

12 Scrub/Shrub,

13 Palustrine Forested Wetland,

14 Palustrine Scrub/Shrub Wetland,

15 Palustrine Emergent Wetland,

16 Estuarine Forested Wetland,

17 Estuarine Scrub/Shrub Wetland,

18 Estuarine Emergent Wetland,

19 Unconsolidated Shore,

20 Bare Land,

21 Open Water,

22 Palustrine Aquatic Bed,

23 Estuarine Aquatic Bed,

24 Tundra,

25 Snow/Ice,

**Appendix III**

1. Deciduous Forest
2. Mixed Forest
3. Evergreen Forest
4. Grass/Pasture
5. Agriculture
6. Palustrine Emergent Wetland
7. Palustrine Scrub/Shrub Wetland
8. Palustrine Forested Wetland
9. Estuarine Emergent Wetland
10. Estuarine Scrub/Shrub Wetland
11. Estuarine Forested Wetland
12. Phragmites
13. Residential
14. Urban
15. Water
16. Bare Land
17. Clouds