Chesapeake Conservancy

EXPLORE.CONNECT.CONSERVE.

Implementing Technology and Precision Conservation in the Chesapeake Bay

By Jeffrey Allenby and Conor Phelan

This report was prepared by the Chesapeake Conservancy and was made possible through financial assistance provided by the Digital Energy and Sustainability Solutions Campaign (DESSC).

For more information about the content in this report please contact the Chesapeake Conservancy at (443) 321-3610 or through our website www.chesapeakeconservancy.org.

The authors would like to extend special thanks to the following people for their contributions to this publication: Stephen Harper, Christopher Hankin, Joel Dunn, Joanna Ogburn, Patrick Noonan, and Paul Hagen.

The Chesapeake Conservancy's mission is to strengthen the connection between people and the watershed, conserve the landscapes and special places that sustain the Chesapeake's unique natural and cultural resources, and encourage the exploration and celebration of the Chesapeake as a national treasure.

The Digital Energy and Sustainability Solutions Campaign (DESSC) brings together information and communications technology (ICT) companies and associations, non-governmental organizations, customers and other stakeholders who recognize the enabling role that ICT plays in improving our environment and driving long-term economic growth

Cover Photo: The image on the cover was created by the author and is a combination of a picture taken from a helicopter and a LIDAR point cloud depicting the same area. While LIDAR data can vary in quality and resolution, even a lower-density dataset, such as this one, can very accurately depict the landscape that it represents, including trees, landcover, and buildings. Photo taken by Jeff Allenby.

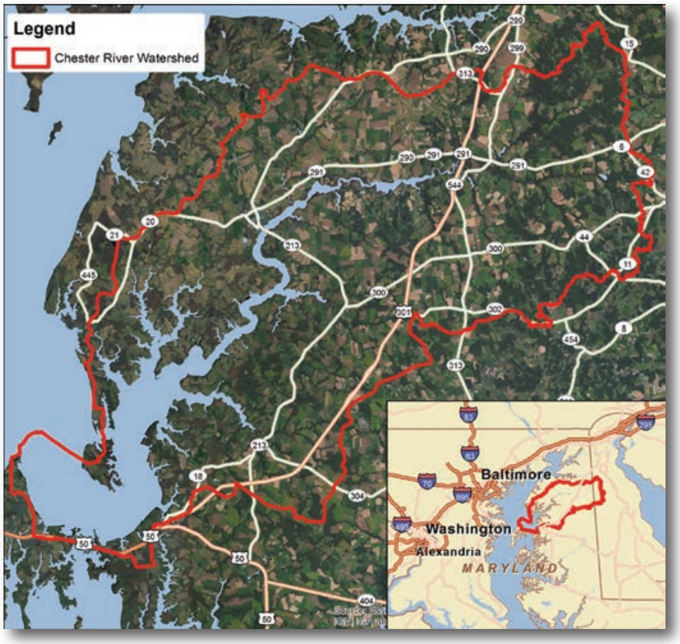
Table of Contents

Executive Summary

Over the last decade, new technologies and more affordable computing power have transformed the way that researchers and conservation managers are approaching decisions of what land to protect. New satellite and aerial collection techniques are generating data that is relevant at the field-scale, allowing conservation and restoration practices to be targeted in the areas where they would have the greatest impact on protecting water quality, priority habitats, and other ecological priorities.

Within the Chesapeake Bay watershed, the establishment of the EPA's Chesapeake Bay Total Maximum Daily Load (TMDL) has placed increased attention on the impact that land use has on water quality. Currently, local, county, and state governments are evaluating available solutions to help meet their requirements for reducing the amount of pollution entering the Chesapeake Bay. As these decisions are made, it will be imperative for conservation organizations to be able to locate, protect, and restore high-functioning landscapes that reduce these nutrient and sediment loads entering the water. Targeting conservation and restoration activites where they will do the most good will help achieve the region's TMDL goals while generating the funding needed to conserve landscapes that provide a suite of ecosystem services in addition to reducing runoff.

In 2012, Chesapeake Conservancy, with the support of the Digital Energy and Sustainability Solutions Campaign (DESSC), explored the role that technology could play in expanding the application of precision conservation; or getting the right practices, at the right scale, at the right time, in the right place. To increase the use of these emerging landscape analysis techniques, DESSC provided the Conservancy with an additional grant to test the most promising technologies and identify how these tools could improve the capabilites of the conservation community, identify the major challenges to implementing them on a wider scale, and better understand the ability for the resulting datasets to highlight priority landscapes within the project watershed.



New tools were used to create high-resolution land use and hydrology datasets in the Chester River watershed and identify areas with an impact on water quality.

For this analysis, Chesapeake Conservancy focused on two main techniques to improve conservation and restoration targeting in the Chester River watershed:

1. Creating a high-resolution land use dataset from aerial imagery, and
2. Mapping concentrated flow paths using LIDAR elevation data.

Using the information from these datasets, a third layer was created to help managers better understand how land use is expected to influence nutrient and sediment loads in waterways. With this information, conservation planning can occur at a landscape scale within the project watershed to idenitfy the stream reaches that will help achieve water quality goals, protect the most important natural landscapes, and direct funding for restoration projects to the areas that will provide the greatest results.

Performing a high resolution landscape analysis for the Chester River watershed was a lengthy, but successful project that created three datasets that will help the Conservancy's partners identify and protect the watershed's highest functioning landscapes. This project has significantly improved the Conservancy's understanding of the intricacies of conducting high-resolution landscape analyses and has provided a strong base from which additional projects will benefit. Over the next year, the Conservancy will distribute the data to local partners, create tools to help organizations with limited technical capabilities improve their conservation targeting, and expand the coverage of high resolution landscape information in other watersheds around the Chesapeake Bay.


Jeff Allenby

Introduction

Within the Chesapeake Bay watershed, the establishment of the EPA's Chesapeake Bay Total Maximum Daily Load (TMDL) has placed increased attention on the impact that land use has on water quality. Local, county, and state governments are currently evaluating various solutions to help meet their TMDL requirements for reducing the amount of nutrient and sediment pollution entering the Chesapeake Bay. As these important decisions are made, it will be imperative for conservation organizations to improve their ability to target high-functioning landscapes and coordinate potential funding sources interested in the multitude of ecosystem services land conservation provides.

Although it has long been accepted that intact riparian ecosystems can prevent a substantial amount of nutrient and sediment pollution from entering waterways, existing datasets used by managers are limited to determining problem areas at the watershed level by looking at which areas have high concentrations of agriculture or impervious surfaces. Many of the existing land use and elevation datasets available to the public lack the resolution needed to identify specific actions that could be taken to reduce the amount of runoff entering the water. Without this more detailed information, and the resulting evidence that land conservation and restoration provide effective solutions to meet TMDL guidelines, many local governments are only focusing on hard engineering solutions such as upgrading wastewater treatment plants and storm water retrofits. While all of these efforts will still help meet the TMDL goals, engineered solutions are typically more expensive than restoration and conservation projects and do not deliver any of the ancillary ecosystem services provided by restoring and protecting natural landscapes.

Over the last decade, new technologies and improvements in affordable computing power have allowed for significant advances in the way that researchers and conservation managers are approaching decisions of what land to protect. Available elevation datasets and satellite and aerial imagery have been steadily increasing in resolution and land managers are gaining access to information at a scale that was previously only available through field monitoring. The ability to conduct a desktop analysis to locate ecosystems that would benefit from conservation or restoration is greatly increasing managers' ability to identify, compare, and prioritize potential projects within an entire landscape. Furthermore, these improvements in the resolution of remotely sensed data are allowing field-scale management decisions to be made about which conservation and restoration practices would have the greatest impact on protecting water quality, priority habitats, and other ecological priorities.

In 2012, Chesapeake Conservancy, with the support of the Digital Energy and Sustainability Solutions Campaign (DESSC), authored a report that explored the role that these emerging technologies could play in precision conservation; or getting the right practices, at the right scale, at the right time, in the right place. This report highlighted five types of landscape analysis that could help improve conservation targeting, the software and data requirements for each, and the potential factors that could either contribute to or hinder their widespread use. The goal was to increase the exposure of some of the most cutting-edge research currently being undertaken throughout the country and to encourage the use of innovative technologies to help improve the effectiveness and efficiency of conservation and restoration programs.

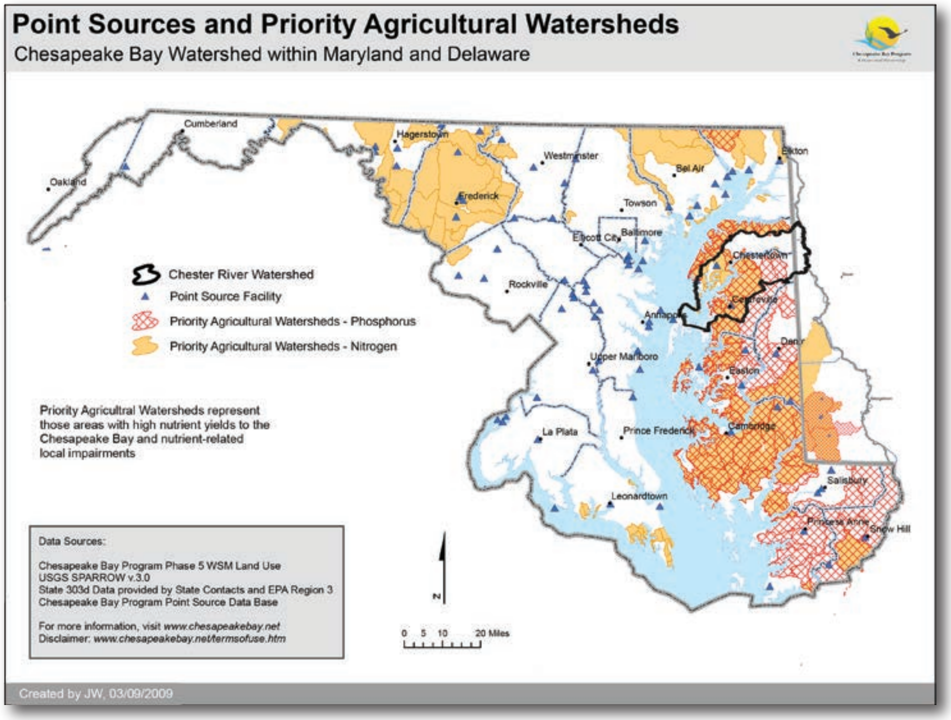


Figure 1: The Chester River watershed has been identified by the EPA as a high priority agricultural watershed for both phosphorus and nitrogen pollution and by the USDA as a showcase watershed for agriculture best management practices. Modified from EPA 2009

As a second phase to this report, DESSC provided the Conservancy with an additional grant to test the most promising technologies identified to better understand the potential uses of these tools within the conservation community, the major challenges to implementing them on a wider scale, and the ability for the resulting datasets to identify priority landscapes within the project watershed.

The Conservancy chose to conduct a watershed-wide landscape analysis in the Chester River watershed to complement their focus on large landscape conservation efforts along the Captain John Smith Chesapeake National Historic Trail and to provide a useful dataset in a watershed that has been identified by the Chesapeake Bay Program as a high priority agricultural watershed for both phosphorus and nitrogen pollution (Figure 1).

For this analysis, the Conservancy focused on two main techniques to improve conservation and restoration targeting in the Chester River watershed:

- Creating a high-resolution land use dataset from aerial imagery, and
- Mapping concentrated flow paths using LIDAR elevation data.

Once these two datasets were generated, they could be combined to create a third dataset that would help managers better understand which stream reaches are expected to have higher than normal levels of nutrient and sediment pollution based on the land that drains into them. With this information, a new level of conservation planning can occur within the project watershed to help achieve water quality goals, protect the most important natural landscapes, and direct funding for restoration projects to the areas that will provide the greatest impact.

The Need for High-resolution Land Use Mapping

Land use and land cover (LULC) data derived from aerial and satellite imagery has been widely available for some time, however the majority of existing datasets are limited by either spatial resolution or by geographic extent. LULC data that were collected and analyzed over large geographies typically have a moderate spatial resolution due to the size of the resulting files. Conversely, datasets with higher spatial resolutions usually cover smaller geographic areas due to the space and computing requirements needed to handle the information. Due to these inherent limitations, large area/low-resolution datasets are often used to characterize the composition of large landscapes, but lack the detail needed to make field-scale management decisions, and small area/high-resolution datasets have the resolution needed to make small scale decisions, but can rarely be used to prioritize between projects within a watershed due to their limited extent.

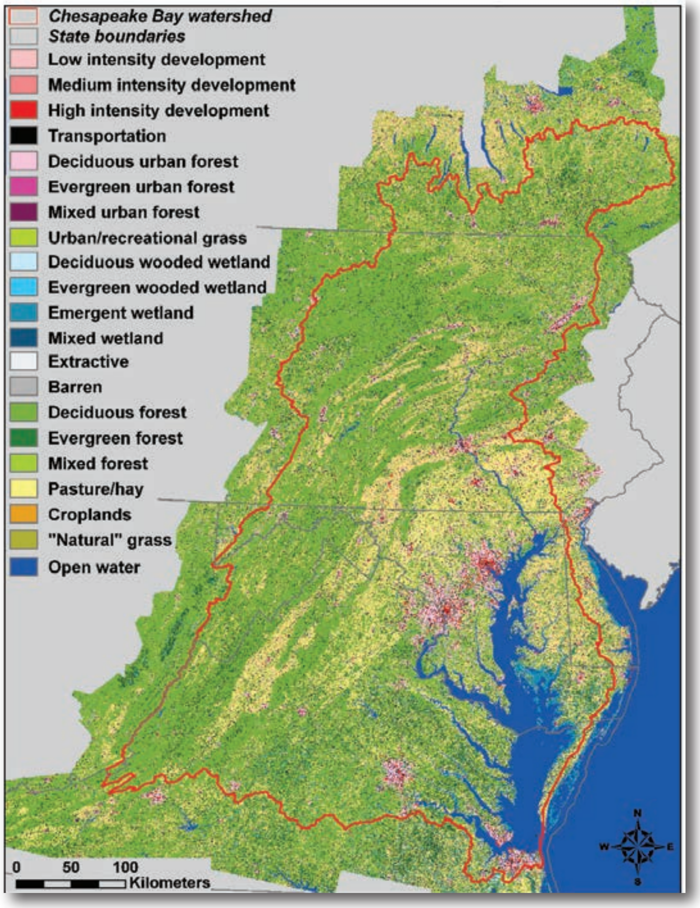


Figure 2: An example of a classified multispectral image depicting the Chesapeake Bay watershed's land cover developed using Landsat 7 TM imagery. Woods Hole Research Center 2000

Since the 1970's, the United States Geologic Survey (USGS) and National Aeronautics and Space Administration (NASA) have been using LANDSAT satellites to determine LULC in the United States. First created in 2001, and subsequently updated every five years, the National Land Cover Dataset (NLCD) is the most comprehensive dataset currently available to researchers and managers and is used by most federal and state agencies, including the Chesapeake Bay Program, to determine land use change over time and to identify priority watersheds that are likely contributing nitrogen, phosphorus, and sediment based on their land use composition. LANDSAT images, and the resulting NLCD data, have a moderate spatial resolution of 30m, which provides a level of detail capable of classifying landscapes across broad areas and has been used successfully for a number of landscape-scale management activities.

The Chesapeake Bay Program has extended the usefulness of this data by creating additional datasets from historical satellite imagery creating a Chesapeake Bay Land Cover Data (CBLCD) series for the years 1984, 1992, 2001, and 2006 that shows how land use has changed throughout the watershed over the last twenty-five years (Irani and Claggett, 2010). Due to its extent and the amount of time it takes to process these large files, however, there is a considerable lag between when the data collection and its release to the public; the current NLCD/CBLCD data used by most federal and state agencies for planning purposes was collected in 2006 but was not released until 2011. As computing power increases and becomes more affordable, this lag will decrease, as evidenced by the 2011 NLCD dataset, which is scheduled to be released in early 2014, almost two years more quickly than the previous iteration, however the usefulness of these datasets for planning purposes is still limited.

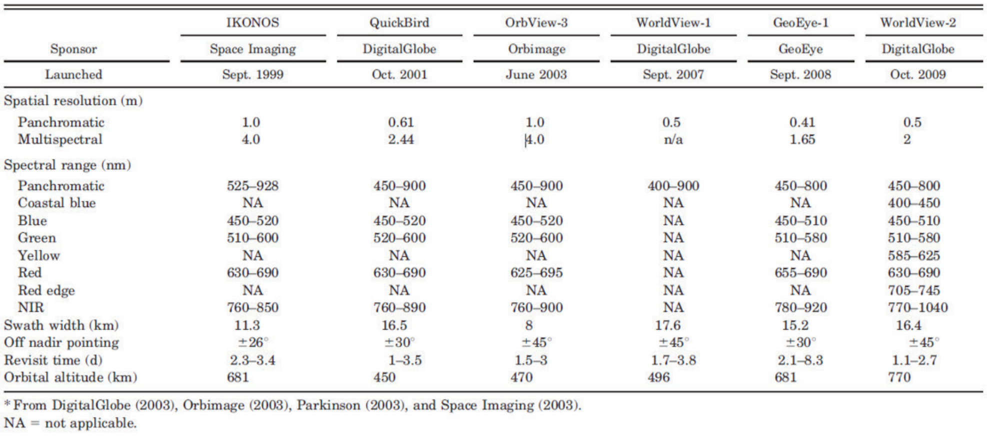
When planners attempt to translate this data into parcel scale management decisions, NLCD and CBLCD data lacks the resolution needed to make informed decisions about where to intercept runoff before it enters waterways. Many agricultural best management practices, such as field-side buffers, ditching, and retention ponds, and some impervious surfaces, including sidewalks and driveways, are often smaller than a single pixel in the LANDSAT image, which means that they are not accurately classified in the datasets used by management agencies and are not accounted for in any models estimating pollutant loads (Claggett et al 2013). In many cases, changes in land use that have occurred since the previous dataset was collected also reduce the relevancy of the information and make it difficult for planners to identify priority landscapes for conservation. Consequently, management agencies, researchers, and conservation organizations are finding that existing datasets are not providing the level of detail they needed to make decisions about what land to protect and what actions would result in the largest reductions of pollution coming off the land.

Recent advances in aerial and satellite imagery interpretation and processing, highlighted in Chesapeake Conservancy's report, provide potential solutions that could help bridge this gap and provide managers with the information needed to target conservation and restoration activities in the areas where they will provide the biggest impact. High-resolution multi-spectral imagery is regularly collected throughout the world by a number of commercial companies with varying spectral and spatial resolutions and costs (Table 1). Commercial satellites range in spatial resolution from 2-10m for multispectral imagery and have between four and eight spectral bands (Klemas 2011).

Commercial aerial imagery tends to have a higher spatial resolution, ranging from 0.3-2m, but lower spectral resolution as most imagery has been collected using either 3-band natural color or color infrared sensors. The prices for both satellite and aerial imagery also vary considerably based on the quality of the data and range from $0.33 to over $30 per km2 depending on if the imagery is archived or if it needs to be acquired. For many conservation organizations and local goverments, the costs associated with commercial imagery are likely higher than budgets can accomodate, reducing the practicality of this type of analysis across a large landscape for conservation purposes.

In the United States, as well as many other countries, aerial imagery is used for a number of different management applications and high quality, publically available, datasets exist. The United States Department of Agriculture's Farm Service Agency acquires aerial imagery during the agricultural growing seasons for the continental United States through its National Agriculture Imagery Program (NAIP). Collected every two years, this data has one meter spatial resolution and 4 spectral bands (red, green, blue, and near infrared) and is freely available through the USGS' National Map (http://viewer.nationalmap.gov/viewer/). For the purposes of large landscape, high resolution LULC analyses, this dataset provides both the spatial and spectral resolution needed to determine unique LULC classes as well as the geographic extent needed to conduct this work on a scale that will make it relevant to planners.

Table 1: Commercial satellite imagery has a variety of costs and resolutions. Aerial imagery can have higher spatial resolution, but often has lower spectral resolution. Adapted from Klemas 2011



	IKONOS	QuickBird	OrbView-3	WorldView-1	GeoEye-1	WorldView-2
Sponsor	Space Imaging	DigitalGlobe	Orbimage	DigitalGlobe	GeoEye	DigitalGlobe
Launched	Sept. 1999	Oct. 2001	June 2003	Sept. 2007	Sept. 2008	Oct. 2009
Spatial resolution (m)						
Panchromatic	1.0	0.61	1.0	0.5	0.41	0.5
Multispectral	4.0	2.44	4.0	n/a	1.65	2
Spectral range (nm)						
Panchromatic	525–928	450–900	450–900	400–900	450–800	450–800
Coastal blue	NA	NA	NA	NA	NA	400–450
Blue	450–520	450–520	450–520	NA	450–510	450–510
Green	510–600	520–600	520–600	NA	510–580	510–580
Yellow	NA	NA	NA	NA	NA	585–625
Red	630–690	630–690	625–695	NA	655–690	630–690
Red edge	NA	NA	NA	NA	NA	705–745
NIR	760–850	760–890	760–900	NA	780–920	770–1040
Swath width (km)	11.3	16.5	8	17.6	15.2	16.4
Off nadir pointing	±26°	±30°	±45°	±45°	±30°	±45°
Revisit time (d)	2.3–3.4	1–3.5	1.5–3	1.7–3.8	2.1–8.3	1.1–2.7
Orbital altitude (km)	681	450	470	496	681	770

* From DigitalGlobe (2003), Orbimage (2003), Parkinson (2003), and Space Imaging (2003).

NA = not applicable.

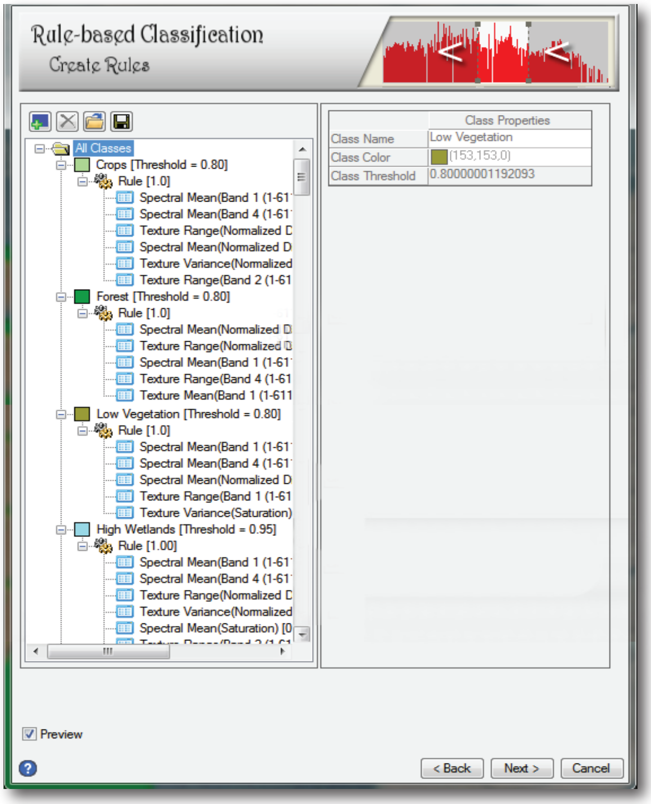
Image Analysis Methods and Findings

Due to a number of factors, Chesapeake Conservancy chose to use the 2011 NAIP imagery dataset, downloaded from the National Map website for its demonstration project in the Chester River. As a result of the size of the downloaded data, the imagery came divided into over fifty segments, each depicting one quarter of a USGS quadrangle. To facilitate the processing of over 12 gigabytes of raw imagery data into LULC data, each quarter quad was analyzed individually using an object-oriented image analysis technique.

The analysis was done in Exelisvis ENVI 5.0 using their rule-based feature extraction workflow. This process uses image segmentation to divide the raw imagery into contiguous zones of visually similar land cover based on each pixel's spectral characteristics. The feature extraction tool uses pattern recognition to identify similar groupings of pixels in the image and then "grows" the initial area to encompass all the surrounding pixels that have similar characteristics. Because the analysis used the rule-based feature extraction workflow, which allows users to define a specific set of spectral and textural "rules" for each LULC class, segments were kept small and the tolerance for merging similar segments was set to a lower level to minimize misclassification of visually similar land use types.

Figure 3: NAIP Imagery was segmented into polygons by grouping pixels with similar spectral characteristics.

This process also makes use of the normalized difference vegetative index (NDVI) to add additional information upon which segments could be separated. NDVI is a comparison of the red and near infrared (NIR) spectral bands and relies on the characteristics of vegetation absorbing red light and reflecting NIR light. This value was especially helpful in the LULC classification due to the limited number of spectral bands in the NAIP imagery and the visual similarity between some of the LULC classes.



- All Classes
  - Crops [Threshold = 0.80]
    - Rule [1.0]
      - Spectral Mean(Band 1 (1-61
      - Spectral Mean(Band 4 (1-61
      - Texture Range(Nomalized D
      - Spectral Mean(Nomalized D
      - Texture Variance(Nomalized
      - Texture Range(Band 2 (1-61
  - Forest [Threshold = 0.80]
    - Rule [1.0]
      - Spectral Mean(Nomalized D
      - Texture Range(Nomalized D
      - Spectral Mean(Band 1 (1-61
      - Texture Range(Band 4 (1-61
      - Texture Mean(Band 1 (1-611
  - Low Vegetation [Threshold = 0.80]
    - Rule [1.0]
      - Spectral Mean(Band 1 (1-61
      - Spectral Mean(Band 4 (1-61
      - Spectral Mean(Nomalized D
      - Texture Range(Band 1 (1-61
      - Texture Variance(Saturation)
  - High Wetlands [Threshold = 0.95]
    - Rule [1.00]
      - Spectral Mean(Band 1 (1-61
      - Spectral Mean(Band 4 (1-61
      - Texture Range(Nomalized D
      - Texture Variance(Nomalized
      - Spectral Mean(Saturation) [0

	Class Properties
Class Name	Low Vegetation
Class Color	(153,153,0)
Class Threshold	0.80000001192093

☑ Preview

Figure 4: Once segmented, a set of rules was created that distinguished the LULC classes from each other.

Once the image was segmented, a set of rules for each LULC class was constructed, setting thresholds for the reflectance value of the four bands of light, the NDVI value, the hue, saturation, and intensity of the natural color image, and the "texture" of these values, which measures how quickly values change in the immediate vicinity of a segment. After the spectral signatures were set for each LULC class, a visual inspection of the image was conducted using a preview of the final classification to determine if any rule was grossly misclassifying land types. Finally, any offending rules were modified and the preview was accurately classifying LULC types, the classification was completed and the output LULC dataset was exported to a raster layer.

One of the benefits of using a rule-based classification, as opposed to an example-based classification where users delineate specific examples for each LULC class, is that the spectral signatures of one image can be imported into the next image after it has been segmented. Due to slight differences between images, the rule set needed to be adjusted to accurately classify the LULC types of each specific image, but the process of classifying subsequent images was drastically shortened by using a base rule set developed for the first image.

Once all of the images were processed into LULC classes, high-quality ancillary data was used to overlay specific land use types that were difficult to discern using the imagery alone. Roads, open water, and wetlands were often misclassified due to slight changes in reflectance, such as sunlight reflecting off the water, or because of their similarity to other classes, such as roads being classified as tilled soil and wetlands being classified as forest. By using these ancillary datasets to classify known land use types, an emphasis was able to be placed on previously unmapped classes, improving the accuracy of their classification. After all of the images were corrected and finalized, they were mosaicked and clipped to the Chester River watershed boundary to create a single watershed-wide LULC dataset.

Finally, an accuracy assessment on the final LULC data was run to quantify how well the analysis classified the imagery. Using a stratified random sampling, 275 sample points were created across all LULC types and the classified values were extracted to the point data. Each point was then manually inspected to determine the actual land use type using the original aerial imagery. Both the classified and actual values were totaled for each LULC class and entered into a matrix to determine the overall accuracy of the dataset (Table 2). Overall, approximately 87% of points were correctly classified, a value greater than the 78% accuracy of the 2006 NLCD dataset (Wickham et al. 2013).

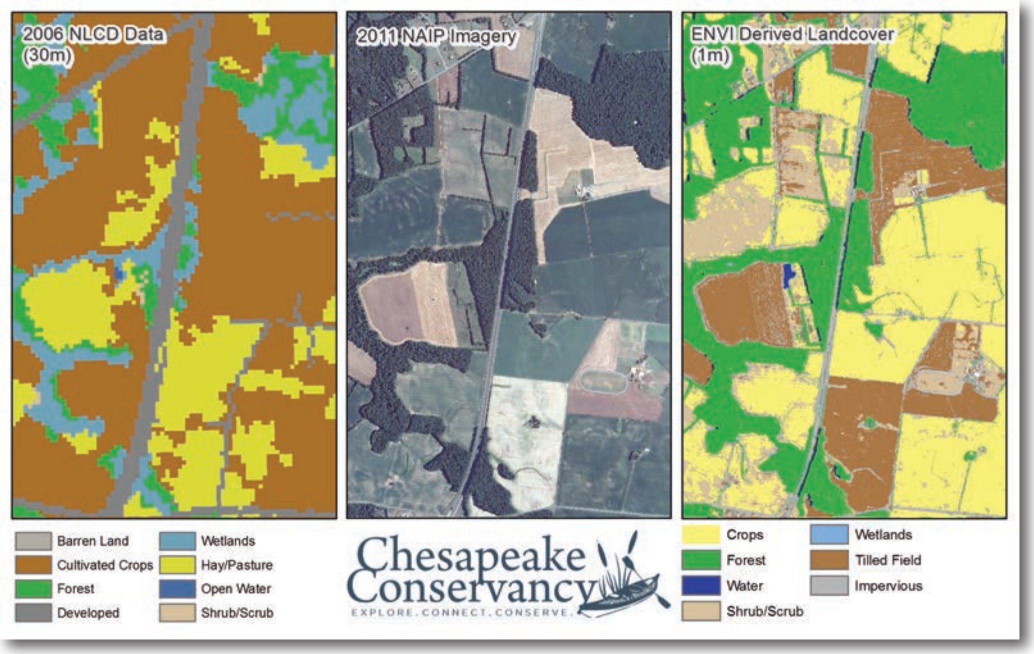


Figure 5: A comparison of the existing NLCD data, the raw NAIP imagery, and the final LULC data created through the rule-based classification.

Table 2: The accuracy assessment shows which LULC classes were classified properly and which were more troublesome. All classes, except impervious, had over an 86% user accuracy indicating a high level of confidence that the landscape was accurately classified.

		Actual							
	Land Class	**Crops**	**Forest**	**Impervious**	**Low Vegetation**	**Tilled Field**	**Water**	**Wetlands**	**Totals**
Classified	**Crops**	45	1		2	1		1	45/50
	Forest	2	45		2			1	45/50
	Impervious	1	2	14	4	2	1	1	14/25
	Low Vegetation		4		43	3			43/50
	Tilled Field			2	2	46			46/50
	Water						24	1	24/25
	Wetlands	1	1			1		22	22/25
	Totals	49	53	16	53	53	25	26	
							Overall Acc.		**86.91%**
							Kappa Statistic		**0.843972**

The Need for Concentrated Flow Path Analysis

Hydrology data in the United States has been collected by the USGS since its inception and stream channels, watershed boundaries, and coastlines that have been mapped in the Chesapeake are available to users through the National Hydrologic Dataset (NHD). This information is the basis for a number of state and federal regulations, including the Clean Water Act, and is critical for protecting important riparian ecosystems and a central component of understanding how land use impacts water quality. Traditionally, the USGS's hydrology data was collected by field observations, using topographic maps, and by manual interpretation of aerial imagery. With the advent of digital elevation models (DEMs) and remotely sensed elevation data, the location of stream channels and other hydrologic boundaries in the NHD have been adjusted using computer modeling to better account for how water is moving across the landscape.

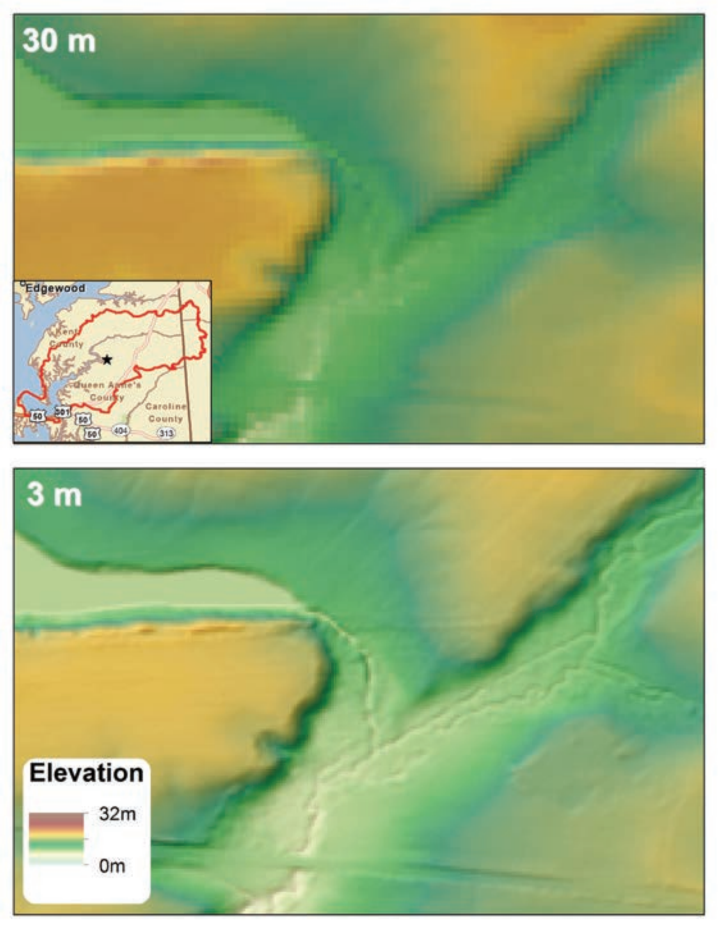


Figure 6: Spatial resolution can have a dramatic effect on the accuracy of stream networks and the features they include. As a result, existing hydrology datasets based on moderate resolution data often do not include smaller features in headwater areas.

Two seamless National Elevation Datasets (NEDs) have been created by the USGS to better model changes in elevation and hydrology; a 30m resolution DEM has been available since the early 2000's and a 10m dataset was completed in the Chesapeake by 2010. Access to the NED datasets have increased the reliability of NHD data and have provided the source information for researchers and modelers to understand how water flows across the landscape, but even the higher resolution NED lacks the resolution to accurately identify smaller features and changes in elevation in areas of low relief.

These datasets are especially limited in the level of detail they contain in upland areas and can miss a large number of headwater streams and field ditches, both of which can deliver large quantities of nutrients and sediment to downstream waterways if they are unprotected. Small streams are often seasonal and are not always apparent in aerial imagery, especially in heavily forested areas, however they can have the largest potential for denitrification when natural landscapes and buffers are intact (Alexander et al 2007). Because they are not included in the NHD, unmapped headwater areas are not protected under most regulations and are vulnerable to development, tilling and other agricultural practices, and degradation.

Before runoff reaches these headwater stream channels, concentrations in water flow are also influencing the effectiveness of agriculture best management practices (BMPs), which, if not accounted for, can significantly impact the amount of nutrients and sediment entering the water. Models determining the effectiveness of filter strips and riparian buffers, as well as regulations and programs that credit them as best management practices, are typically based on the assumption that water flows evenly across the landscape and interacts with buffers equally at all points. In nature, this is rarely the case and certain areas will receive more runoff than others, which can quickly overwhelm the filtering capacity of these BMPs. By designing filter strips and riparian buffers to accommodate the water flowing off the landscape, variable width buffers can provide almost twice the cost-efficiency as traditional buffers by providing increased ecosystem services such as water quality improvement, erosion control, and wildlife habitat protection (Qiu and Dosskey 2012).

Using only the NED and NHD data, Managers interested in designing variable width buffers or protecting headwater streams currently do not have access to the information they need to make fully informed decisions. While the NHD data provides accurate locations of known stream channels and watershed boundaries, this information does little to help landowners identify high quality functioning landscapes or problem areas that would benefit from protection or restoration. Similarly, 30m NED data lacks the resolution to identify hydrology accurately and even 10m data has trouble detecting the minute changes in elevation that influence how water reaches a stream channel. The coverage of Light Detection and Ranging (LIDAR) elevation data has expanded significantly over the last decade and provides managers and planners with high-resolution elevation data in these areas that can be used to address both of these issues (Figure 7).

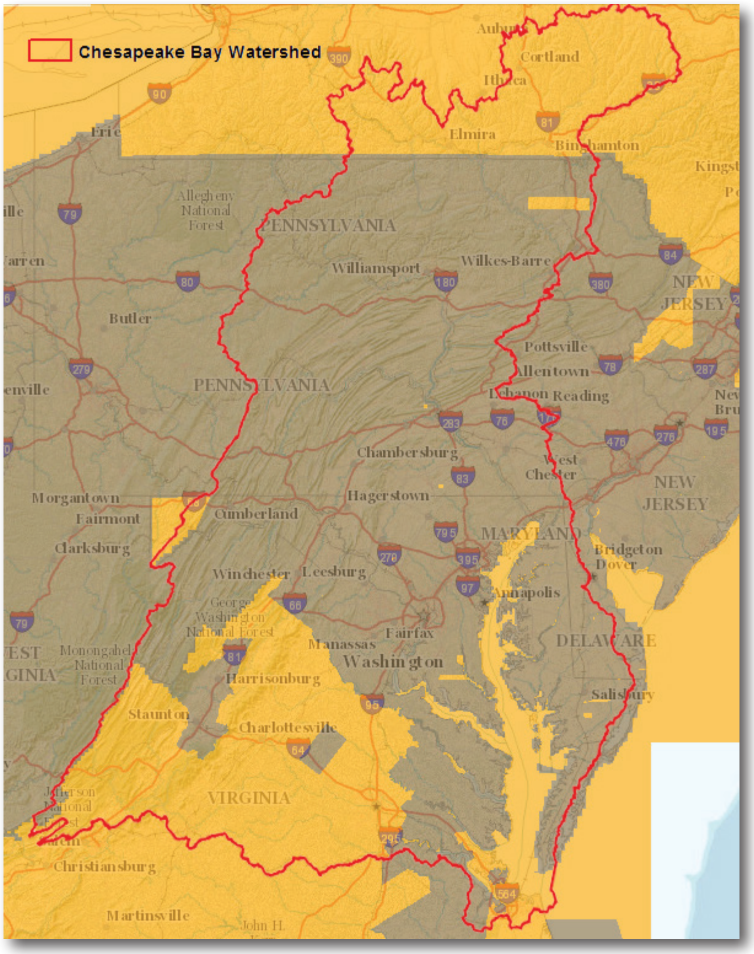


Figure 7: Availability of NED data throughout the Chesapeake Bay watershed. Gray areas have 3m elevation data available and yellow areas have 10m data.

LIDAR data varies in spatial resolution based on the specifications to which it was collected, but most datasets have 1-3m horizontal resolution with a vertical accuracy of better than 37cm. Many coastal areas of the United States have been mapped with LIDAR for the purpose of mapping floodplains and there is complete coverage of the Chesapeake Bay watershed in Maryland, Delaware, Pennsylvania, and New York and partial coverage in Virginia (Figure 7). This data is available from each state, typically with 2m resolution, or through the National Map as a 3m DEM. For the purposes of large landscape, high resolution concentrated flow path mapping, these LIDAR datasets provide the spatial resolution needed to determine complex hydrology in headwater areas as well as the geographic extent needed to conduct this work on a scale that will make it relevant to planners.

Concentrated Flow Path Analysis Methods and Findings

Chesapeake Conservancy used the USGS's 3m DEM for this project because it was already processed into a raster image and would not be impacted by the Maryland/Delaware border transecting the watershed boundary, avoiding any latent artifacts resulting from each state contracting and processing the raw LIDAR point data to different specifications.

Again, due to the size of the elevation dataset, the LIDAR raster data was downloaded as a series of 48 tiles that covered the project area. Because hydrology is dependent on having complete information within a watershed, all of the tiles were mosaicked together to create a single elevation raster for the entire Chester River watershed, which was then clipped to the watershed boundary to remove any extraneous data.

The analysis was conducted in ESRI ArcGIS 10.2 using the TauDEM 5.1.1 ArcGIS toolbox, a free extension written by Dr. David Tarboton at Utah State University available at http://hydrology.usu.edu/taudem/taudem5/.

Although ArcGIS's Spatial Analyst extension has native hydrology tools, they rely on a D-8 analysis while TauDEM allows for either a D-8 or D-infinity analysis. The D-infinity analysis is more processor intensive and typically takes longer, but provides a more accurate representation of how water flows across the land, especially in low-relief areas like the coastal plain. (For information on the difference between D-8 and D-infinity analyses, please read Chesapeake Conservancy's previous report, available at: chesapeakeconservancy.org/Reports.)

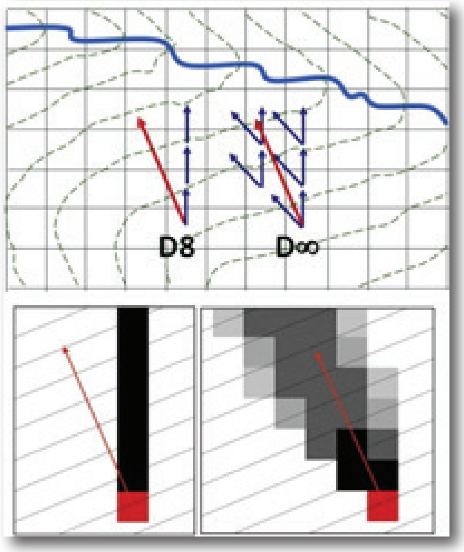


Figure 8: D∞ models provide a more accurate representation of flow across low relief landscapes.
Tesfa et al. 2011

The process was started by performing a "pit remove" function on the elevation data. Pits are typically single pixels in DEMs that are completely surrounded by higher terrain and are artifacts in the data that interfere with the routing of flow across DEMs. They are corrected by raising their elevation to the minimum elevation of the pixels around them, removing their impact on the hydrology of the area. This process, while not necessary, helps create a more complete hydrology in the watershed by avoiding breaks in flow paths and stream channels where there is not supposed to be one.

After the base elevation dataset was corrected, the "D-infinity flow direction" function was used to map how water is flowing between cells. This analysis maps the flow direction of each cell by determining how water is distributed to the two adjacent cells with the largest change in elevation. Once this dataset was created, the "D-Infinity contributing area" was calculated to measure how many upland cells are draining into each cell. This process uses the flow directions to map the contributing areas and is useful for determining where water is concentrating before entering a stream.

To extract the location of these concentrated flow paths, a conditional statement was used in ArcGIS's Spatial Analyst toolbox to extract all values above 250 to determine only the areas that have at least 250 m^2 (2690.98 ft^2) draining into them (Figure 9). Contributing area is often used to determine the start of headwater stream channels, however this threshold is highly variable and dependent on a number of other characteristics including soil type, slope, and convexity of the drainage area. Although the value of 250 m^2 does not necessarily signify the start of a headwater channel, it functions well to separate concentrated flow paths in upland areas from background data representing areas where sheet flow is likely occurring.

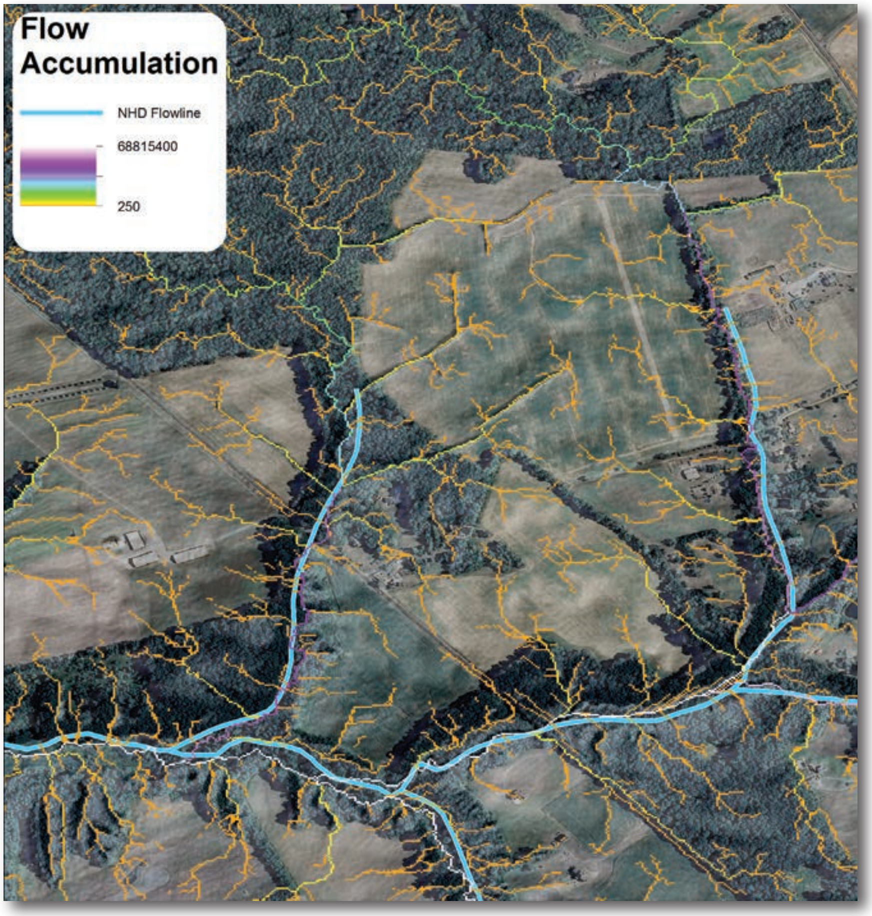


Figure 9: Extending hydrology analyses into headwater areas can provide a much greater understanding of how water is flowing of the land and where it could be intercepted before reaching water ways. Using these artificial concentrated flow paths also provides a much higher level of detail than the existing NHD flow lines.

Combining LULC and Hydrology Information to target Best Management Practices

High resolution LULC and hydrology data are powerful tools on their own and can help inform a large number of management activities, but when they are combined, they can provide managers with an in-depth understanding of how these landscape features are interacting with each other to impact water quality and identify areas that are likely to be problems. While high-resolution data will not completely replace field verification, it will allow organizations to use this data as a primary targeting tool and subsequently spot check various areas to ensure that the information is accurate. Remotely targeting, and later monitoring, conservation and restoration efforts has the potential to significantly reduce the cost and time involved with field monitoring and lessen resource constraints experienced by many conservation organizations.

There is currently no landscape-scale information available to managers to help target restoration efforts at anything less than the sub-watershed level. The Chesapeake Bay Program and state agencies have done a good job of identifying impaired watersheds using the Chesapeake Bay Watershed Model and are directing restoration efforts and funding to the watersheds that need the most help, but because field-scale information is lacking, these efforts may not be effecting the greatest benefit possible. With this increased knowledge of how water is moving off the land, funding through the USDA's programs to implement agricultural best management practices can also be prioritized to areas where they will make the biggest difference in water quality.

Figure 10: Poorly planned agriculture can result in significant nutrient and sediment loads entering the Chesapeake Bay. While often easy to spot visually, identifying problem areas such as this one over a large landscape requires innovative methods that make use of the best data available.

Many conservation and restoration programs in the Chesapeake Bay have come under scrutiny for not achieving the reductions in sediment and nutrient pollution that were expected, but recent research shows that as little as 15% of farms could be contributing a majority of the nutrient and sediment pollution into waterways (Wisconsin Buffer Initiative 2005). If best management practices are not targeted towards these problem areas, the reductions that are needed, and expected, will not be realized and the funding will have been spent in less than optimal places.

Traditionally, to gain this level of understanding, very small scale projects were conducted to analyze and understand how an individual property, or small group of properties, could be modified to improve the water quality coming off the land. These studies typically involve a significant amount of effort and field measurements, making them a costly endeavor for most landowners. Furthermore, these small-scale analyses are almost always conducted "on-demand" after a landowner has made the decision to install a BMP or restore a natural ecosystem on their property. In this regard, the resulting information is providing managers with an understanding of how to properly design a project for that particular landscape, but not how it might compare to other potential projects or where funding would best be spent.

Conducting a high-resolution landscape analysis on a watershed scale will give managers the ability to determine which areas should be priorities for conservation and restoration as well as a better understanding of what actions would best address the unique issues associated with each particular project. Having this information from the start allows conservation organizations and state agencies to target outreach and education efforts to the landowners in priority areas and allows them to help property owners understand what actions would have the greatest impact from the start, making the entire process easier to embrace. For conservation organizations, having more advanced targeting data will also drastically improve their ability to obtain grants for conservation and restoration projects as they will be able to include the expected impacts in applications, something many grant giving organizations look for when reviewing potential projects.

Landscape Analysis Methods and Findings

Chesapeake Conservancy used a weighted contributing area analysis to better understand how land use and hydrology interact with each other in the Chester River watershed. TauDEM's "D-infinity contributing area" function includes the opportunity to input a weight grid that specifies the contribution to flow for each cell. This analysis will help identify stream reaches that have high concentrations of agriculture and impervious surfaces and which are well protected by forests and wetlands.

The first step to conduct this analysis was to reclassify the LULC data to represent the relative potential for nutrient and sediment pollution. Impervious surfaces, tilled soils, and agriculture were given a higher weight while low vegetation, wetlands, and forests were given a low weight. Water was not given a value as it did not contribute any sediment or nutrients and likely provides a reduction in these values through natural processes.

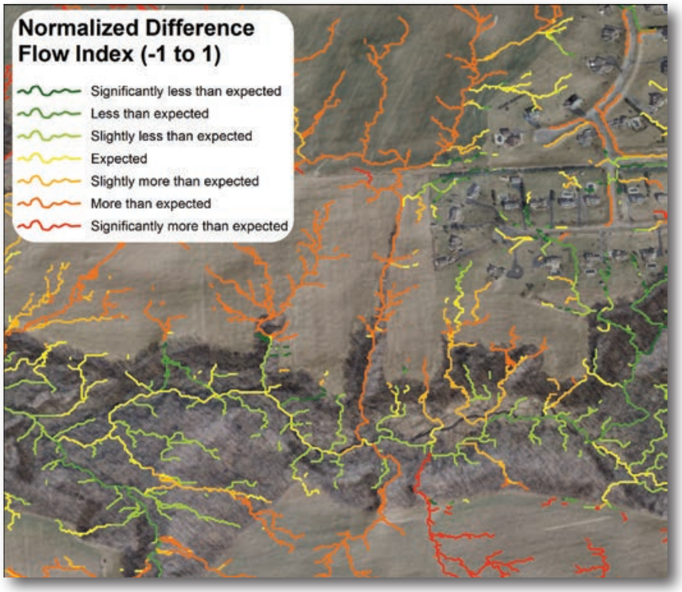


Figure 11: NDFI flow paths can help managers identify problem areas by highlighting where larger than expected flows are entering main waterways.

After the LULC data was reclassified, the data had to be resampled to match the spatial resolution and extent of the elevation dataset for processing purposes. A majority filter was used for the resampling, meaning that each new pixel, which encompassed nine of the original dataset's pixels, took on the value of the majority of the sub-pixels. In many cases, this removed small errors in classification and the accuracy of the resampled LULC dataset improved to 88.3%. To match the extents of the two datasets, the resampled LULC layer was masked to the elevation dataset and then the new LULC layer was used as the weight grid for the "D-infinity contributing area" function. The output of this analysis provided a weighted flow accumulation that could be compared to the unweighted accumulation layer to better understand where increased or decreased flows could be expected based on the land use composition of their contributing areas. Borrowing from the concept of NDVI, staff created a "normalized difference flow index," or NDFI, that compares the weighted and unweighted flow accumulations to understand how land use impacts nutrient and sediment loads using the equation:

$$NDFI = \frac{(Weighted\ flow\ accumulation\ -\ Unweighted\ flow\ accumlation)}{(Weighted\ flow\ accumulation\ +\ Unweighted\ flow\ accumulation)}$$

This equation creates a raster layer with values between -1 and 1 that depicts how far above or below the expected flow the weighted flow is; values closer to 1 represent areas that have much higher than expected nutrient and sediment loads and values closer to -1 represent areas that have much lower than expected nutrient and sediment loads.

This information will be extremely helpful to managers because it not only shows where concentrated flow paths are, it also helps prioritize areas that have the largest need for conservation and restoration projects. Planners can use the NDFI values to understand which landscapes are doing a better job of reducing pollutant loads before they enter the water, most likely due to natural ecosystems, and which areas are likely underperforming due to high concentrations of impervious surfaces and agriculture.

Lessons Learned

Like many projects investigating new techniques and software, this analysis had a number of issues that required a significant amount of troubleshooting. Additionally, every step of the analysis took considerably longer than expected due to the size of the imagery and elevation datasets and there were a number of subtle issues with data management that required very specific solutions that were not immediately apparent. For organizations considering this type of analysis, it is recommended that workstations have at least a four-core processor and eight gigabytes of memory. The Conservancy has invested in computers capable of handling larger datasets, with four-core processors and 16GB of RAM, but processing time was still the largest component of the process and something that should be considered for organizations considering expanding this type of analysis to a larger geography.

Lessons learned from LULC analysis

One of the biggest decisions associated with the LULC analysis was which imagery to use. There is a variety of high-resolution commercial imagery available, and in general, the more bands of imagery there are, the easier it will be to separate different types of land use. After getting estimates for the cost of these datasets, however, it became apparent that this would be a prohibitive factor if this type of analysis were to be completed throughout the entire Chesapeake Bay watershed.

NAIP imagery provided an appealing solution as it is collected regularly, allowing for land use change analyses, and was freely available to everyone. A major early concern was the ability to separate LULC classes using only four imagery bands, as opposed to LANDSAT's 7 bands, however, for the purposes of the project it was more than adequate, especially when supplemented with ancillary datasets for some of the more complex land use types.

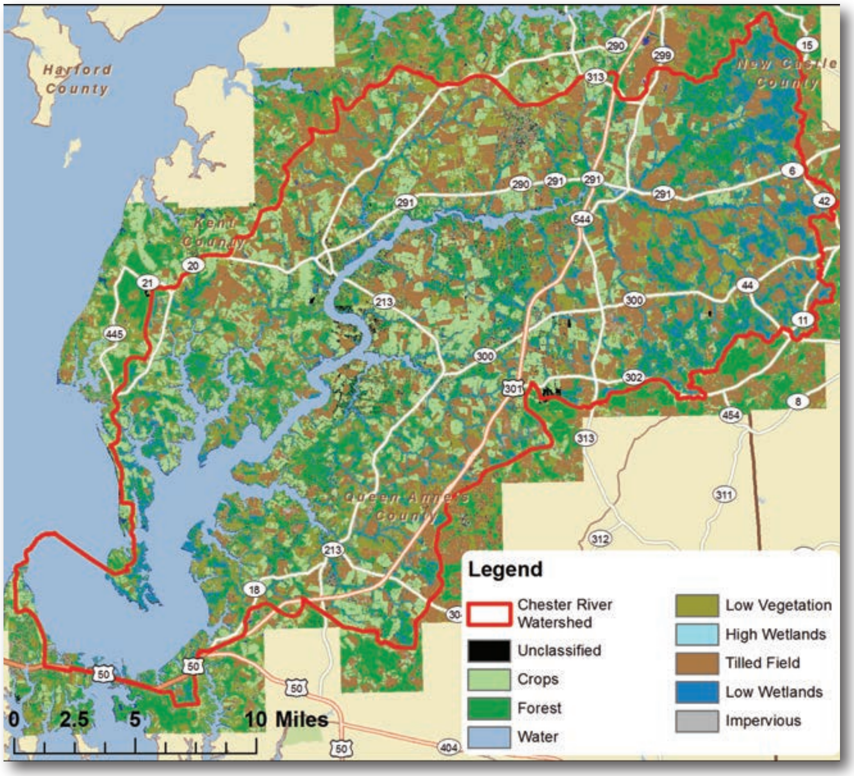


Figure 12: Final LULC dataset for the Chester River watershed.

If this analysis were to be undertaken for regulatory purposes, or with government support, it is recommended that the USDA's Natural Resource Conservation Service (NRCS) be included as they have access to detailed information for the majority of farms in the region. With this data, an analysis would be able to further segment the "agriculture" class by crop type and farming practices, which will permit a more detailed analysis that incorporates nutrient applications and crop efficiencies. Currently, these datasets are not publicly available, making it extremely difficult to incorporate this information without NRCS' involvement.

From the start, it was apparent that classifying the entire watershed at once was not going to be feasible, so the decision was made to process imagery tiles one at a time. This improved the rate at which images could be classified, however it was still a lengthy process. The initial segmentation of an image tile took on average 3-4 hours and each adjustment in the merge and edge segmentation settings before finalizing the segmentation took 5-10 seconds to display in the preview window, creating a potentially lengthy process. Once the segmentation settings was adjusted and finalized, it took another 1-2 hours to calculate the spectral characteristics for each segment before work could begin creating the rules for each class. This process increased in efficiency over time as edge segmentation and merge settings required fewer adjustments, however the time to segment the image and calculate the spectral characteristics remained fairly consistent. The speed of these processes would likely increase with a more powerful processor or increased memory in the computer, something Chesapeake Conservancy plans to invest in for future analyses.

The use of a rule-based feature extraction was also beneficial for a number of reasons. First and foremost, as mentioned previously, a rule-set created for one image can be exported and reused for additional images that are collected in the same time period. Creating the first rule-set was time consuming and required both patience and an understanding of how different land use types reflect and absorb certain bands of light, however this effort drastically reduced the time it took to process subsequent images. Again, every time the ranges on a rule were adjusted, there was about a 10 second delay before the preview window was updated. On average, each LULC class had 4-8 rules and each rule required a substantial amount of fine-tuning to ensure it was not under or over classifying the target class. As a result, even with the base rule-set it could still take close to an hour of adjustments before an image could be classified.

Second, the rule-based classification was a better choice than the example-based classification because as you select more examples, the time it takes to process the spectral characteristics increases dramatically. This was especially evident in more texturally complex landscapes, such as forests, due to the relatively small size of segments and the variability in spectral characteristics. In one case, after selecting a small section of forest that included about 500 segments, it took close to five minutes before the next selection could be made and each subsequent selection took almost as long. The time it took to build a robust set of examples, and the fact that they could not be reused on other images, made the rule-based classification a significantly better choice for a large landscape.

Lessons learned from concentrated flow path analysis

The hydrology analysis was more straightforward than the LULC classification and the process of identifying concentrated flow paths was better established through previous research. One of the largest decisions was whether to use LIDAR data from each state or to download the USGS's 1/9th" DEM. Both provided significant improvements over the existing datasets, however the USGS's data was chosen for the reasons stated previously. Downloading data from the National Map was extremely easy and mosaicking and clipping the datasets provided few issues.

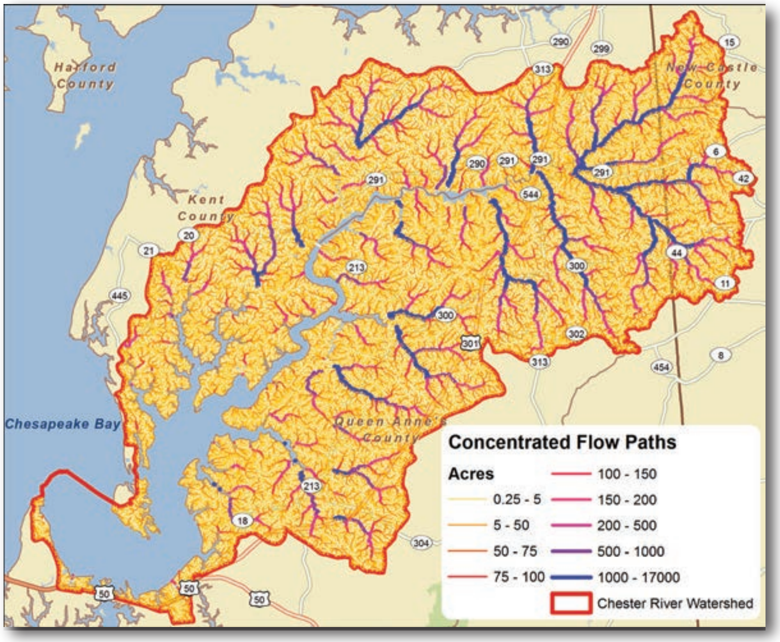


Figure 12: Final CFP dataset for the Chester River watershed showing the number of upland acres draining into each pixel.

The process for installing the TauDEM toolbox is much improved with the most recent update to TauDEM's software and staff had no issues installing it onto a new workstation during the project. Originally mentioned in Chesapeake Conservancy's previous report, installing TauDEM does require an administrator to be logged onto the computer, which could be an issue for some organizations, but will not be an issue for most users.

Two smaller issues were encountered while conducting the hydrology analysis. The first was unavoidable and a result of the base LIDAR datasets being collected at different times. In Maryland, the state partnered with counties to collect the LIDAR data, and as a result, different counties were done at different times. Kent

Island and a small portion of the lower Chester River were collected in 2003, as part of Anne Arundel and Queen Anne's County, however the rest of the Chester River in Kent and Queen Anne's County was flown in 2006. During the processing of these datasets, there was no overlap left at the edges, resulting in a thin line about 16m wide of no data that disrupted the hydrology throughout this gap. The data between Maryland and Delaware did not have this issue, despite being collected at different times, as an overlap was left during the processing.

The second issue encountered was a matter of processing that is easily avoidable once it is known. During the initial unweighted flow accumulation calculation, TauDEM's "Calculate contributing areas" tool creates a "specific catchment area" unit that is calculated as the number of cells multiplied by the grid cell length, or cell area divided by cell length. The original projection for the DEM was a geographic coordinate system, which uses degrees as the horizontal unit, and when the specific catchment area was calculated, it was a fraction of the output unit that the weighted analysis would use. Staff needed to re-project the base DEM raster to a projected coordinate system, which uses meters as the horizontal unit, and rerun both the flow direction analysis and the contributing area analysis for the units to be the same as the weighted analysis. Once this was completed, there were few other issues with the hydrology analysis. Overall, the hydrology took approximately 4-5 hours to complete the flow direction and the contributing area analyses.

Lessons learned from the NDFI analysis

As the last stage, most of the initial issues of this project had been worked out. Deciding on the best way to combine the two datasets took some research and analysis to choose what level of detail was needed for the project purposes. While increasingly complex models may have provided a more explicit output detailing the finite quantity of nutrients or sediment that were expected to come off of the land, the results of these models are typically scrutinized by both conservationists and land owners as "not being accurate enough" for planning or regulatory purposes. The decision was made that the project would be better served by creating a targeting layer that identified hot-spots within a watershed that could be investigated in person.

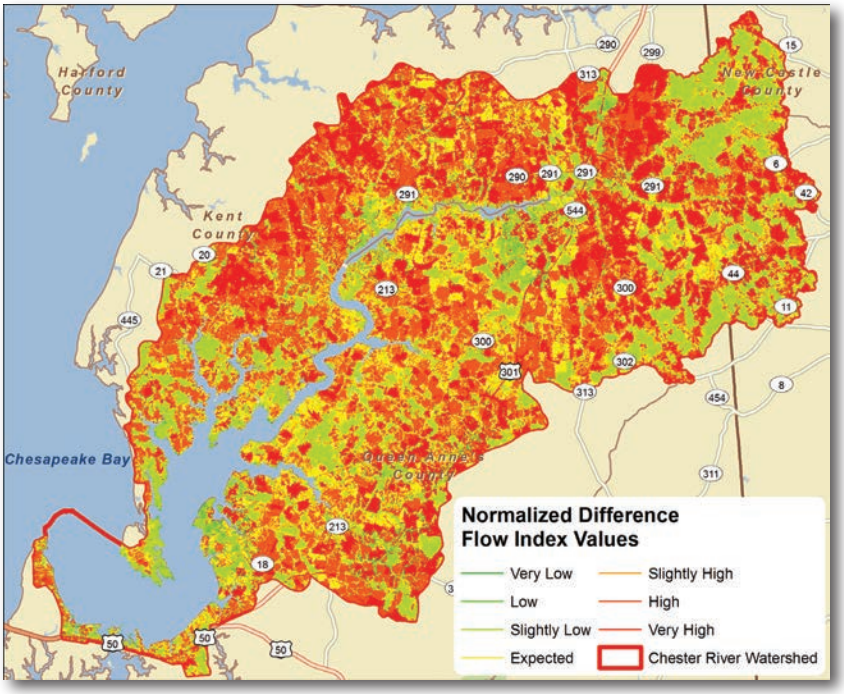


Figure 12: Final CFP dataset for the Chester River watershed showing the Normalized Difference Flow Index (NDFI) of stream reaches.

As a result, the goal for the final step in the analysis was to identify areas that had a large amount of impervious surfaces and agriculture that did not interact with many natural ecosystems before entering the water. Using a normalized difference equation was selected to address this issue as it created an easy to understand scale showing both conservation priorities and restoration priorities.

The projection issue described in the last section was only discovered during the weighted flow analysis and it took a bit of troubleshooting to figure out how far back the analysis needed to go to create a valid dataset. After this issue was solved, the weighted analysis' units matched the unweighted result, providing the necessary datasets for the final calculation. Once the unweighted and weighted flow accumulation layers had been created successfully, it was a simple raster calculation to calculate the NDFI. Overall the weighted contributing area analysis and the NDFI calculation took about 10 hours to complete.

Recommendations

Performing a high resolution landscape analysis for the Chester River watershed was a lengthy, but ultimately successful process that resulted in a dataset that will be useful for a number of planning purposes. The information generated from this project will be tremendously helpful to the Conservancy's partners allowing them to identify and protect the watershed's highest functioning landscapes. More importantly, having these datasets will help these organizations successfully apply for funding to restore some of the region's most threatened landscapes.

This project has significantly improved Chesapeake Conservancy's understanding of the intricacies of conducting high-resolution landscape analyses and has provided a strong base from which additional projects will benefit. As the Conservancy works to expand the coverage of high-resolution information and promote the benefits of these improved datasets, the process of converting raw imagery and elevation data to useful products will become increasingly streamlined and benefit a greater number of organizations. Chesapeake Conservancy has identified a few areas where additional efforts could increase the awareness and usefulness of technology and precision conservation throughout the Chesapeake Bay watershed.

Distributing data to partner organizations

One of the challenges that must be overcome first, however, is how to distribute the resulting datasets to organizations in a form that they can manage. Many smaller organizations have limited GIS capabilities and if they have the necessary software, often times their computers will not be able to handle larger datasets. the Conservancy is working to identify potential solutions to help partners access the high-resolution data efficiently through both physical and web-based solutions.

ArcGIS offers a free data viewer called ArcGIS Explorer that will allow users to view the final datasets without purchasing a full copy of ArcGIS. To facilitate the viewing of data on slower computers, the Conservancy has segmented the final LULC, hydrology and NDFI datasets into smaller datasets representing the eleven HUC-12 sub-watersheds that compose the Chester River watershed. Each sub-watershed contains 20-76 square miles of data, which should be more manageable for most organizations with limited computing power. The Conservancy will distribute DVD's with the data to partners in the coming months and will continue to offer the data to new partners as needed.

Over the next year, Chesapeake Conservancy also hopes to create web-based mapping applications containing the information that will make these datasets available to the public and to groups without GIS capabilities. These tools will allow everyone to explore the information and better understand how their land may or may not be impacting water quality in the Chester River. Additionally, a web-based toolbox is planned that will allow users to virtually experiment with various best management practices to better understand how they will impact water flow off the land and what solutions offer the best combination of reductions and cost.

Expanding the coverage of High-resolution data in the Chesapeake Bay

Not every organization will have the technical capabilities or software and hardware needed to accomplish an analysis such as the one completed in this project, however the level of information that can be generated will transform what the conservation community is able to do and effectively direct staff and financial resources where they are needed most. While conservation organizations can continue to function successfully without more detailed analyses, conservation and restoration actions may not generate the outcomes hoped for and likely will not maximize the use of limited resources.

Moving forward, there are two main directions that should be pursued to expand the creation and use of high-resolution landscape data throughout the Chesapeake Bay. When possible, Chesapeake Conservancy should work to help other organizations obtain the necessary skills and tools needed to undertake analyses in their own project areas. In other cases, organizations will not have the ability to make these commitments and will need to work with the Conservancy and other partners who do have these capabilities to conduct projects in their focus areas. Each option will have its own costs and benefits and it will take a combination of the various opportunities to improve conservation targeting in the region.

Figure 13: The final datasets were segmented into eleven sub-watersheds to make it easier for partners to access and use the data efficiently.

Conducting analyses in-house

Organizations that have the resources to devote to these projects will need to make a commitment to invest in the hardware, software, and training that are needed to efficiently conduct an advanced landscape analysis. Chesapeake Conservancy found that running the routine on a moderately powerful computer was inefficient and typically used all of the computer's resources, making it unusable for other tasks while processing imagery or hydrology. If an organization operates throughout a large area, or plans on making landscape analysis a core part of their operations, investing in a powerful workstation and training staff on the various software packages may make sense. Groups interested n pursuing this route can expect to spend at least $1,500 to purchase a higher powered workstation and non-profit organizations will be able to purchase single user licenses of ENVI and ArcGIS for around $600 total, although local governments and other for-profit groups will have to pay considerably higher for the programs.

Once the hardware and software have been acquired, training for ENVI and ArcGIS is available for little to no cost; Exelsvis provides complimentary access to all of its multi-day training sessions for non-profit users and there is a large online user base for ArcGIS that provides instruction for most issues users may come across. Most of the existing training materials are not focused on conservation targeting, however, and it would be helpful if future trainings were offered with a focus on applying these technologies specifically to identify protection and restoration priorities.

Although it will be a significant investment, high-resolution landscape data can provide a significant advantage to conservation organizations by increasing their capacity to target conservation and restoration projects in places that matter and their ability to justify projects when applying for grants. Ultimately, however, whether or not it makes sense to invest in the capabilities to conduct an analysis will depend on how much an organization will use this information and how much the data will improve their return on investment.

Partnering with other organizations to conduct analyses

If conducting high-resolution landscape analyses will likely be a small part of an organization's programs, the cost of getting the software and hardware and training staff may not be the best investment of time and resources. For many organizations, it will make more sense to partner with another organization who has the capabilities to conduct an analysis in their project area. Additionally, there are often multiple organizations and local governments concerned with conservation and restoration in each watershed, opening the possibility for cost-sharing agreements as well as coordinated planning at a landscape scale.

Chesapeake Conservancy has made high-resolution, landscape-scale analysis a central part of its operations throughout the watershed and is actively interested in connecting with other organizations to provide the Conservancy's capabilities to partners who are interested in having the data, but not investing in the infrastructure themselves. Through its Conservation Innovation Center, the Conservancy is helping other groups understand what can be done with the higher resolution data and developing additional projects in high-priority watersheds. Resulting from the work in the Chester River, the Conservancy was able to partner with the National Oceanic and Atmospheric Administration's Chesapeake Bay Office to conduct a high resolution landscape analysis in the Choptank River watershed. The resulting data from both of these projects will be shared with local governments and conservation organizations to help target restoration and conservation efforts that will maximize the water quality benefits in each watershed.

Engaging Colleges and Universities

Partnering with universities and colleges also presents an attractive option for expanding the use of advanced spatial analysis techniques for organizations with limited resources. In many cases, higher education institutions are open to partnering with non-profit organizations to provide services that will help their students gain useful experience and training in new techniques.

During the summer of 2013, Chesapeake Conservancy was able to partner with the University of Richmond to conduct an analysis of four sub-watersheds in the James River watershed. These watersheds were identified as priorities by the Conservancy's Envision the James community-based conservation planning process and were studied to better understand riparian buffer quality and to detect opportunities for restoration projects. Through this partnership, Conservancy staff trained a group of students to use the ENVI software to create a high-resolution land use dataset, moderate resolution hydrology, due to a lack of LIDAR data, and a modified NDFI layer. This information is now being used to help with planning efforts and the students were able to gain much needed experience that will make them more attractive when applying for jobs after graduation. After a successful completion of this project, University of Richmond professors have expanded the program to other areas of the James River watershed.

Figure 14: Chesapeake Conservancy staff teach University of Richmond students how to conduct a landscape analyis on the James River in Virginia to help identify priority buffer restoration projects.

Chesapeake Conservancy is also exploring partnerships with other universities throughout the watershed in an effort to expand the coverage of high resolution datasets. There has already been significant interest from a number of schools and Conservancy staff hope that as these programs expand, local organizations will be able to access the datasets they need to make informed decisions. Chesapeake Conservancy is exploring additional opportunities to match schools with local organizations who have a need for high-resolution datasets but do not have the capacity to generate it themselves and sees this type of partnership as one of the most important drivers of the expansion of precision conservation throughout the Chesapeake Bay watershed.

References

Alexander, Richard B., Elizabeth W. Boyer, Richard A. Smith, Gregory E. Schwarz, and Richard B. Moore (2007). “The Role of Headwater Streams in Downstream Water Quality” Journal of the American Water Resources Association 43(1): 41-59

Claggett, Peter R., Frederick M. Irani, and Renee L. Thompson (2013). “Estimating the Extent of Impervious Surfaces and Turf Grass Across Large Regions.” Journal of the American Water Resources Association (JAWRA) 1-21

Irani, F.M, Claggett, P.R. (2010). “Chesapeake Bay watershed land cover data series.” U.S. Geological Survey Data Series 2010-505

Klemas, V., (2011). “Remote sensing of wetlands: case studies comparing practical techniques.” Journal of Coastal Research, 27(3): 418–427

Qiu, Z. and M.G. Dosskey (2012). “Multiple Function Benefit – Cost Comparison of Conservation Buffer Placement Strategies.” Landscape and Urban Planning 107(2): 89-99.

Tesfa, Teklu K., David G. Tarboton, Daniel W. Watson, Kimberly A.T. Schreuders, Matthew E. Baker, Robert M. Wallace (2011) “Extraction of hydrological proximity measures from DEMs using parallel processing” Environmental Modelling & Software 26(12): 1696-1709

Wickham, James D., Stephen V.Stehman, Leila Gass, Jon Dewitz, Joyce A.Fry, and Timothy G. Wade, (2013). “Accuracy assessment of NLCD 2006 land cover and impervious surface.” Remote Sensing of Environment, 130: 294 – 304

Wisconsin Buffer Initiative, (2005). “A Report to the Natural Resources Board of the Wisconsin Department of Natural Resources by the University of Wisconsin-Madison, College of Agricultural and Life Sciences.” Accessed at: http://faculty.nelson.wisc.edu/nowak/reports/nrbFinalReport.pdf.

Chesapeake Conservancy
EXPLORE. CONNECT. CONSERVE.

716 Giddings Avenue, Suite 42
Annapolis, Maryland 21401
info@chesapeakeconservancy.org
443-321-3610