Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Dataset Lessons Learned and Stakeholder Outreach

Cassandra Pallai

Geospatial Program Manager, Chesapeake Conservancy

Margaret Markham

Geospatial Analyst, Chesapeake Conservancy (former)

Kathryn Wesson

Geospatial Analyst, Chesapeake Conservancy (former)

Louis Keddell

Geospatial Analyst, Chesapeake Conservancy

Jarlath O'Neil-Dunne

Director, University of Vermont Spatial Analysis Laboratory

Anna Royar

GIS Technical Lead, University of Vermont Spatial Analysis Laboratory

Peter Claggett

Research Geographer, United States Geological Survey

Executive Summary

Part 1: Lessons Learned from the Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project

To model, track, and manage statewide progress toward reaching water quality goals, the Chesapeake Bay Program Partnership (CBPP)—a regional partnership that helps lead, direct, and manage restoration efforts in the Chesapeake Bay—relies heavily on land use and land cover data. Before 2015, CBPP calculated watershed states' pollution contribution numbers using the 30-meter resolution National Land Cover Database (NLCD). However, NLCD often fails to match the accuracy and resolution of local datasets being employed by state and county planners. This has resulted in inconsistencies when comparing local evaluations with CBPP's modeling, which has made planning, implementation, and reporting very difficult.

CBPP recognized that having a high-resolution land cover dataset for the entire watershed and surrounding counties would not only improve their ability to quantify the effects of existing and in-progress restoration, but also would have very real implications on planning new restoration projects. Thus, in 2015, CBPP commissioned a high-resolution land cover dataset for the Watershed. Three main contractors, the Chesapeake Conservancy (CC), the University of Vermont Spatial Analysis Laboratory (UVM), and Esri partner WorldView Solutions, Inc. (WVS), created a one-meter-resolution dataset for the Chesapeake Bay Watershed and all the counties that intersect its boundaries. Within the project geography, CC classified Maryland, New York, Washington, D.C., and West Virginia; UVM was responsible for Delaware and Pennsylvania; and WVS covered Virginia (Figure 1). This groundbreaking product covers over 100,000 square miles, and has a resolution of one meter.

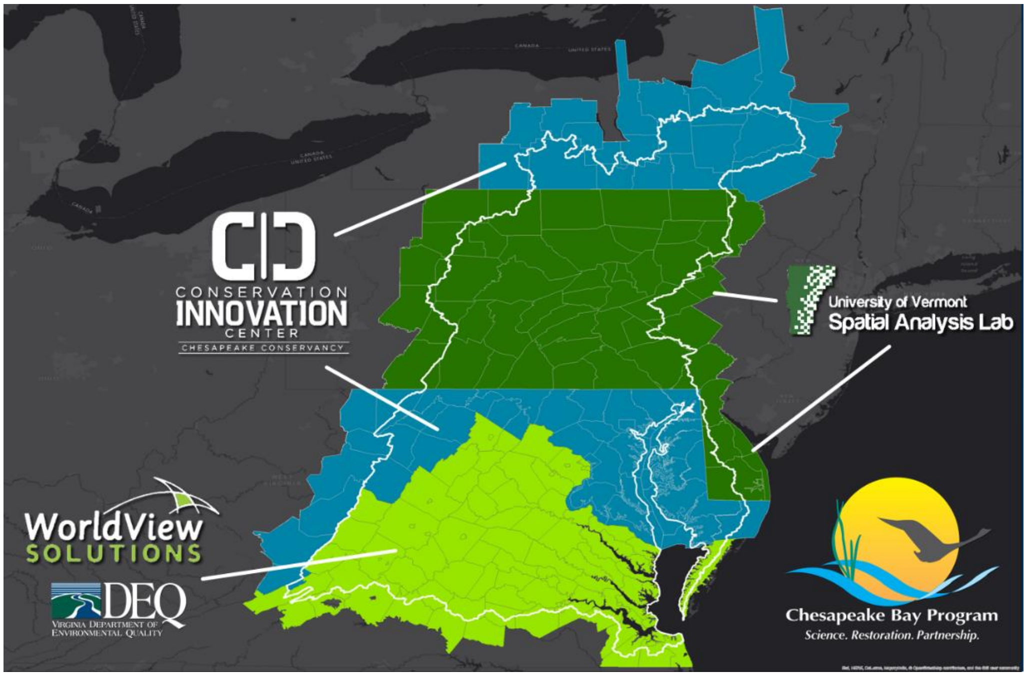

Figure 1. Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project contractor mapping extents.

Part 1: Lessons Learned from the Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project details the successes, challenges, and recommendations for major components of the project design including: classification methodology; source data planning; data review and county requirement gathering; and setting accuracy and deliverable expectations. Only feedback from CC, UVM, and CBPP has been included; this document is not representative of WVS.

Classification methodologies differed between CC, UVM, and WVS. These differences stemmed from both expertise and available source data. CC and UVM used a rule-based object-oriented approach, which was applied to one-meter resolution National Agriculture Imagery Program (NAIP) imagery as well as orthophotos to classify features. Additional imagery and data, such as regional vector data, was used to further aid in the classification, and final corrections were done manually with extracted imagery segments. These contractors favored this mixed-data strategy, which incorporated local, state, and federal vector data directly into the land cover in locations where they met quality standards. Depending on quality, data either wholly substituted for a class (e.g. structures) or was edited and then augmented remote sensing of that class.

Depending on local capacity, counties and/or state GIS staff were contacted for their feedback about the local input data and the land cover classification; most conversations were had one-on-one via email or over phone. Comments for CC were submitted through customized ArcGIS Web Apps. An innovative protocol also was devised to tackle accuracy assessment issues that arose when applying traditional methodologies to the high-resolution land cover classification. This new process ensured a fair assessment that could be translated to future iterations of high-resolution classification projects, and was thoroughly tested to prove its transferrable nature. The results of the accuracy assessment averaged 90% for the entire watershed when measured on a statewide basis. This was met with CBPP approval given the context of unforeseen complications throughout the project, such as fluctuating feature definitions, which compelled CBPP to lower their initial class-by-class accuracy expectations. CC's public website hosts the final land cover classification and accuracy assessment methodology for any stakeholder to access the information. Each of the project design components provided successes and challenges. Based on these, CC has recommendations for future iterations of the land cover project.

The classification methodology produced many successes and challenges, and further steps are recommended to streamline the process. First, project definitions and standards should be set prior to a project's beginning. Most importantly, class definitions and minimum mapping units must be established prior to the start of a project. The time spent on corrections should be distributed according to classes that are particularly important, such as impervious features. In the event that a feature definition is edited during the project, one person should be responsible for correcting the misclassifications in the sections that have been completed under the original definition. A pilot study of hands-on classification time/error ratio should be done, as should an initial assessment of different classification software platforms for their performance in specific project areas.

Future classifications also should start with a comprehensive data management strategy. Pre-project time should be spent standardizing and troubleshooting data preparation tools and procedures. From the start of the project, all data products should be projected to a single prearranged coordinate system, and snapped to a uniform raster. A classified 100-meter buffer should be included around each deliverable, and seamlines should be checked between imagery tiles and political boundaries throughout the process. One analyst should be assigned to each data collection geography (e.g. individual counties or cities), and if leaf-on and leaf-off imagery will be processed, the same person should be responsible for generating classifications from both. If planimetric datasets meet quality standards, those features should be incorporated into the final product as part of the post-processing system.

Given the lack of a corrections interface and the numerous issues encountered by CC with ENVI, eCognition is recommended as a classification software as it provides a multitude of more advanced processing capabilities and is generally more adept at classifying imagery in regards to accuracy, aesthetics, and time. If ENVI is selected for its upfront investment savings compared to eCognition, several recommendations should be followed. Vegetation indices such as the Normalized Difference Vegetation Index (NDVI) should be calculated through the classification software. Any datasets, such as Lidar, that are below a predetermined usability threshold should not be merged with the imagery to create the classification segments. Temporary files produced by ENVI's preview window should be regularly deleted using programs such as WizTree, and the attribute data option when exporting data should be unchecked to prevent extremely large file outputs. Based on preliminary tests, and testimonial from UVM, a layered classification approach within eCognition is recommended as more efficient as opposed to classifying all classes simultaneously. Deep learning or ArcGIS's image processing capabilities with raster analytics and distributed processing may also provide a viable and time-saving alternative to either of these platforms.

A mixed-source data strategy, which incorporates local, state, and federal vector data into the land cover classification workflow, should be used to reduce processing time and to satisfy local stakeholders' demands for incorporation of their existing land use information (e.g. impervious surface extent). Collection of component data, particularly imagery and elevation data, should be coordinated across the entire classification area in advance of the project start. The quality and speed of classification would be improved significantly with uniform, high-quality, high-resolution imagery that has at least four bands (red, green, blue, and near-infrared), and Lidar or photon Lidar Digital Surface Model (DSM) that matches the collection timeframe of the imagery. For example, Hexagon Geospatial 60-centimeter resolution imagery would cost approximately $9/square mile, plus $2.25 for an auto-correlated photon DSM. It is on a two-year collection schedule. If a coordinated collection of imagery, Lidar, and/or planimetric data is not possible, time should be invested in planning a product delivery schedule around the sources of component data and any pending data updates or collection dates. Additionally, all component data should be assembled as far ahead of the project start date as possible, and a data storage plan should be developed with a standard naming convention to improve organization, including a template that organizes temporary files from final files in a central document with staff assignments. If possible, a specialized team should take responsibility for obtaining and preparing all raw data in order to ensure a reliably consistent product to classify. Data quality guidelines and usability standards should be applied to the land cover classes and states separately instead of one uniform data development time standard, depending on component data quality and availability.

While the land cover review process provides the opportunity to develop relationships with contacts at counties across the watershed, it needs to be planned and budgeted as a significant project component. Targeted engagement is needed with the counties about a regulatory land cover project, both in advance of initiation and during data creation, so as to reduce gaps in communication. Project leaders should share with counties or other stakeholders, well in advance, any expectations for involvement, as well as land cover samples and a realistic review schedule. Another possible strategy is to designate one dedicated staff person to manage all county outreach and engagement through the duration of the project to ensure that adequate attention is paid to the process and to avoid bottlenecks with otherwise-engaged project staff.

Accuracy standards and review methods also should be decided at the beginning of a project. Expectations for accuracy should be set and shared widely well ahead of the project start, and be based on the quality of imagery and component data, as well as the project timeframe. An assessment protocol needs to be agreed upon with CBPP and other stakeholders before the project start that defines the number of accuracy assessment samples to be gathered and the appropriate area of accuracy analysis, preferably by state, and any other relevant methods considerations. CC recommends an assessment that is object-based rather than pixel-based, that utilizes the feedback of multiple reviewers, and that incorporates each of the component datasets such as imagery and vector information that have been implemented in the land cover classification. A fuzzy accuracy assessment is the most ideal methodology for this type of data, and should be conducted impartially without looking at the classification. Samples are to be based on the different land cover classes and should not intersect with transition zones from one class to another.

Ideally, future land cover classification projects should try to avoid rolling deliverable timelines as they are difficult to maintain and track, and instead plan to send all data drafts out to clients at the same time. Additionally, CBPP should be involved in initial contact with counties to further facilitate communication regarding project details. This includes coordinating ahead of time in regard to contacting counties and assisting with introductions, making sure everyone is on the same page, and explaining the purpose of the project, how it will benefit users, and deliverable appearance expectations.

Part 2: Feedback from Stakeholders about the Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project

After the conclusion of the Chesapeake Bay high-resolution land cover project, CC spearheaded an outreach campaign to amass existing and emerging use cases, gauge usability, and collect feedback about valuable resources and edits for any updates to come. The outreach sought to gather feedback from users and their thoughts on the review process, data resolution, and classification categories. *Part 2: Feedback from Stakeholders about the Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project* details the results of this outreach campaign.

Through email, over 300 people were contacted, including individuals from federal, state, and local levels. Contacts were gathered from CBPP, high-resolution land cover webinars, and existing partnerships. Of those contacted, 36 responded to surveys, and approximately 26 individuals from varying fields of study and geographic location were verbally interviewed. Though representation was lower for Washington, D.C., New York, and West Virginia, overall a fairly diverse and representative set of participants was engaged across the watershed.

Most respondents were excited about the high-resolution data, with the wetlands, tree canopy, and impervious surface classes being the most popular. While users were eager to utilize it, there were a number of obstacles listed for why users had a difficult time analyzing the land cover, both generally and specifically to CBPP's land cover product. For general high-resolution products, users cited lack of staff time, staff training, available processing power, and financial capacity to incorporate high-resolution land cover into their work. Many also mentioned their work being very project specific, often not including need for high-resolution data. Some users listed similar reasons for hampering their ability to use CBPP's high-resolution land cover, but also listed reasons more specific to region and class. For those users that were able to download and start analyses, some expressed concerns with class consistency across state lines, and the varying resolutions and times of year inherent in the data as a result of the mixed-methods approach. Others listed misclassified shadows and specks in forest classes as problematic, and the accuracy of the wetland and shrubland classes as not being adequate for their purposes. The partial coverage of land cover for New York and West Virginia also prohibited some users from working with land cover in those states, as they required statewide coverage for their projects. There also was some concern with understanding the accuracy of the dataset, and the need for updates to the data at consistent intervals was expressed.

Users who were familiar with the coarser-resolution NLCD were excited about the new availability of the high-resolution land cover, but there were obstacles specific to translating classes between the two datasets, and incorporating the new data into models and projects traditionally configured for NLCD. To make data translation easier, respondents replied that there was no need to mimic every single land cover class in NLCD, but identified the wetland, forest, and agricultural classes as being high priority in their work and ideal classes to mimic in future updates. The majority of respondents preferred aggregating the data up to a 10-meter resolution for the sake of processing time, and suggested a compromise: classify land cover at the one-meter resolution, and package the publicly-available data at 10-meter resolution.

Recommendations that came out of the outreach campaign should be taken into consideration, but with the understanding that the CBPP high-resolution land cover has been available to the public for less than one year. As a result, the pool of users polled was fairly small and many were not yet able to evaluate the data, much less incorporate it into their current work. Users would like to see wetlands and tree canopy classes separated into subgroups like those in NLCD's classification (i.e. woody wetland, emergent wetland, deciduous forest, evergreen forest, etc.). Other suggestions included a more robust shrubland class, and converting from raster to vector format for some users. New York and West Virginia users would like to see expansion of land cover boundaries for state-wide analysis, and the inclusion of an outcrop and mine class for West Virginia. Many respondents would like to see more outreach and engagement prior to the rollout of a new dataset, as many interviewees reported being confused about the details and hosting locations of the CBPP land cover dataset and CBPP land use dataset. Confusion and need for clarification between the CBPP land cover and land use datasets were pervasive issues throughout the outreach process. Moreover, those who were aware of the differences and had listened to the webinar regarding methodology indicated that a similar webinar focused solely on example uses of land use data and how to manipulate the CBPP dataset would be particularly helpful.

Table of Contents

Part 1: Lessons Learned from the Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project

Introduction

In 2010, the Environmental Protection Agency (EPA) put the Chesapeake Bay on a pollution diet to restore the health of the estuary. The diet, called a Total Maximum Daily Load (TMDL), strives to decrease the amounts of nitrogen, phosphorus, and sediment in the water by 20 percent by 2025. The TMDL has directly affected state and countywide policies and regulations throughout the watershed and stimulated a number of new pollution reduction initiatives at local levels.

To model, track, and manage watershed-wide progress toward reaching these water goals, Chesapeake Bay Program Partnership (CBPP)—a regional partnership that helps lead, direct, and manage restoration efforts in the Chesapeake Bay—relies heavily on land-use and land-cover data. Before 2015, CBPP calculated its pollution metrics using the 30-meter resolution National Land Cover Database (NLCD). However, NLCD often fails to match the accuracy and resolution of local datasets being employed by state and county planners. This has resulted in inconsistencies when comparing local evaluations with CBPP's modeling, which has made planning, implementation, and reporting very difficult.

CBPP recognized that having a high-resolution land cover dataset of the entire watershed would not only improve their ability to quantify the effects of existing and in-progress restoration efforts, but also would have very real implications on planning new restoration projects. Thus, in 2015, CBPP commissioned an upgrade the land cover data for the entire watershed. Three main contractors, the Chesapeake Conservancy (CC), the University of Vermont Spatial Analysis Laboratory (UVM), and Esri partner WorldView Solutions, Inc. (WVS), created a land cover dataset with one-meter pixel resolution for the Chesapeake Bay Watershed and all the counties that intersect its boundaries. Within the project geography, CC classified Maryland, New York, Washington, D.C., and West Virginia; UVM was responsible for Delaware and Pennsylvania; and WVS covered Virginia. This groundbreaking product covers over 100,000 square miles.

The Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project consisted of six general phases: 1) establishing data standards, 2) data preparation, 3) automated feature extraction, 4) manual corrections, 5) data compilation, and 6) stakeholder feedback. The following is a set of after-action summaries for each one of these phases along with recommendations on how to carry the next high-resolution land cover mapping project for the Bay. In particular, this document details the successes, challenges, and recommendations for these major components of the project including: classification methodology; source data planning; data review and county requirement gathering; and setting accuracy and deliverable expectations. Only feedback from CC, UVM, and CBPP has been included; this document is not representative of WVS.

Twelve classes were included in the areas mapped by CC and UVM. General class definitions are as follows, and limitations to these definitions are discussed throughout the document. For more specific definitions, please see the project's Image Classification Key.

Land Cover Classes for Delaware, Maryland, New York, Pennsylvania, Washington, D.C., and West Virginia

Water	All areas of open water, generally with less than 25% cover of vegetation/land cover. This includes water-filled backyard pools, ponds, lakes, rivers, natural tidal pools in wetland areas, and boats that are not attached to docks. MMU = 25 square meters, approximately 5 meters wide
Wetlands	Low vegetation (see *Low Vegetation* description) areas that intersect or are near National Wetlands Inventory (NWI) layers (Estuarine and Marine Wetland, Freshwater and Emergent Wetland, Freshwater Forested/Shrub Wetland, and Riverine), that are visually confirmed to have wetland characteristics (i.e. a look of saturated ground surrounding the vegetation), and that are located along major waterways (i.e. rivers, ocean). Areas of low vegetation near the NWI layers are included if they are visually confirmed to be wetland ecosystems. Woody vegetation is excluded from this class. MMU = 225 square meters
Tree Canopy	Deciduous and evergreen woody vegetation of either natural succession or human planting that is over approximately 5 meters in height. Stand-alone individuals, discrete clumps, and interlocking individuals are included. MMU = 9 square meters
Shrubland	Area of both/either deciduous and evergreen woody vegetation that is between approximately 2 and 5 meters in height. Stand-alone individuals, discrete clumps, and interlocking individuals are included, as are true shrubs, young trees, and trees or shrubs that are small or stunted because of environmental conditions. MMU = 225 square meters
Low Vegetation	Low vegetation is plant material of natural succession or human planting that is less than approximately 2 meters in height. This includes visibly tilled fields (with or without vegetation), lawns, nursery plantings with or without tarp cover, and natural ground cover. MMU = 9 square meters
Barren	Areas void of vegetation consisting of natural earthen material regardless of how it has been cleared. This includes beaches, mud flats, dirt roads, and bare ground in construction sites. MMU = 25 square meters
Structures	Human-constructed objects made of impervious materials that are greater than approximately 2 meters in height. Houses, malls, and electrical towers are examples of structures. MMU = 9 square meters
Impervious Surfaces	Human-constructed surfaces through which water cannot penetrate, and that are below approximately 2 meters in height. This includes asphalt, concrete, gravel, pavement, treated lumber (e.g. docks and decks), etc. MMU = 9 square meters, minimum 2 meters wide for linear features
Impervious Roads	Impervious surfaces that are used and maintained for transportation, as denoted by planimetric data. MMU = 9 square meters, wider than 2 meters
Tree Canopy Over Structures	Tree canopy that overlaps with sections of impervious surfaces that have been classified as structures. MMU = 9 square meters
Tree Canopy Over Impervious Surfaces	Tree canopy that overlaps with sections of impervious surfaces that are not roads or structures, including sidewalks and parking lots. MMU = 9 square meters
Tree Canopy Over Impervious Roads	Tree canopy that overlaps with sections of impervious surfaces that have been classified as roads. MMU = 9 square meters

From these mapping efforts has resulted one of the most detailed and highest-quality land cover products ever produced in the United States. This product must be maintained and updated on a regular basis if it is to be of value for its intended goals. Given the investment made in producing a watershed wide line cover data set one approach to carrying out annual updates would be to focus on areas that have undergone significant change. New technologies could help to identify those areas that have changed the thus updates could be focused on these areas. This would require high-resolution imagery, but such data could be obtained from federal state or local sources at a reasonable for no cost. Watershed-wide land cover updates should occur every five to 10 years, synchronized with new imagery and Lidar collects.

Classification methodology

Methods used for classification, definitions, and software.

Because of the variability of the data across the various jurisdictions, no single automated routine could be developed to extract land cover. Instead, classification methodology differed between CC, UVM, and WVS based on expertise, resources, and data availability. A generalized workflow of expert remote sensing systems was employed and then modified for each area based on the availability of the source datasets and the properties of those datasets. CC and UVM used a rule-based object-oriented approach, and one-meter resolution National Agriculture Imagery Program (NAIP) imagery as well as orthophotos to classify features. Additional imagery and data was used to further aid in the classification, and final corrections were done manually using the extracted imagery segments. WVS implemented an example-based methodology with Virginia orthophotos and ancillary data. CC's detailed methods are available for download here.

Definitions

Class	Features that have been categorized based on their appearance and/or land use.
Client	CBPP, the counties, or any third-party involved in the review process or who has tried to use the final land cover classification for their own projects.
Dataset	A collection of related Geographic Information System (GIS) data.
Definition query	Selecting a subset of features from a dataset, based on particular attributes, using a tool through ArcMap.
Deliverable	General term for any portion of the land cover classification that went out to a CC client for review or as the final product.
eCognition	Image analysis software used to classify features.
ENVI	Image analysis software used to classify features.
Esri	Company who owns ArcMap GIS software.
Example-based classification	A classification scheme that uses a series of analyst-defined feature samples as examples of rules by which to categorize the remaining features in an image.

Floating point data	Information associated to points in a dataset that is not relative to the point locations.
Geodatabase	Esri method of storing GIS data.
Manual corrections	Method by which analysts hand-edit corrections to the classification through ArcMap.
Merge	Combining multiple datasets, such as stacking different types of imagery into one composite image.
Minimum mapping unit	The finest resolution at which a classification category is actively mapped and corrected.
Model	Visual geospatial processing workflow that can be customized to user needs.
ModelBuilder	An application in ArcMap that allows a user to create and edit models which automate sequences of analyses for more efficient geospatial processing.
Planimetric data	Two-dimensional geographic representations of features, such as roads and water bodies.
Remote sensing	Capturing data from satellite or aircraft by scanning the surface of the earth instead of physically observing it.
Ruleset	A set of attributes, such as texture or area, where the user dictates the value range in order to define a feature.
Rule-based classification	A classification scheme that uses a series of “if-then” statements dictated by the analyst to categorize features in an image based on its pixel values.
Segments	Groups of pixels that define objects in an image.
Snapping	Aligning a raster dataset’s pixels to a particular extent.
Post-processing	Refining classified data to prepare it for final delivery.
Product	Any portion of the land cover classification that was delivered to a client.
QA/QC	A process for managing data, which stands for *Quality Assessment/Quality Control*.
Usability threshold	A standard of quality of a particular dataset for its ability to be applied to a project.
Versioned editing	Allows users to edit the same dataset simultaneously without corrupting it.

Successes

In general, costs were kept low by using a blend of contractors and subcontractors. This worked well given the constrained timeframe of the project, in that it enabled portions of the classification to be delegated to various organizations rather than solely relying on a single contractor to complete the entire project. A rule-based classification approach was another success; it permitted analysts to develop a classification ruleset that they could apply to other locations. Generally, it outperformed an example-based classification approach, which tended to be more time consuming in choosing the samples, particularly because samples could not be recycled across sets of imagery. This approach required an upfront training investment, but worked well once analysts were comfortable with creating rulesets.

Various combinations of datasets worked well in identifying different classes, e.g. intensity helped in delineating impervious surfaces and water. Additionally, using planimetrics to help delineate impervious features improved efficiency and, for the most part, accuracy of the results.

Manual corrections was a crucial step in the derivation of land cover products. While time-consuming and costly, manual interpretation was the only way to ensure a high level of quality. Given the fixed budget and a large number of technicians working on this process, the main challenge was ensuring standards were met without exceeding the available monetary resources. Having technicians that were properly trained along with the management structure that carried out periodic reviews and quality checks proved to be instrumental to ensuring consistency. Although advances in artificial intelligence and new data sources will likely reduce the need for manual interpretation of future products these developments are unlikely to eliminate the need altogether.

Challenges

One of the challenges of having multiple teams working on a single project was ensuring consistency. Project data standards, to include image interpretation keys, were established at the initiation of the project, but only between CC and UVM as WVS was contracted separately by the State of Virginia Department of Environmental Quality. This was a far from an ideal situation as it resulted in one contractor of three mapping to different specifications.

Another major challenge presented to the classification workflow was the lack of client-established land cover class definitions and minimum mapping units. CC and UVM made every effort to determine a classification schema that as mutually exclusive and totally exhaustive, with clear and concise definitions and minimum mapping units. However, class definitions and standards evolved throughout the project, especially in the first six months of classification work, often in response to challenges that threatened the timeline of the project, or to developments in the thinking of the CBPP Land Use Work Group. This created a situation in which certain areas had to be re-mapped, a less than efficient use of resources. For example, in initial class definitions, barren features were defined as regions devoid of vegetation and consisting of natural earthen material regardless of how they had been cleared. Challenges arose in determining if sparsely vegetated features such as construction sites, landfills, and mining regions fit the initial definition laid out for the classification of barren features. Spectral similarity, particularly in problematic imagery, between impervious surfaces and barren features further contributed to these issues. Under revision and recommendation, the definition for barren was updated to accommodate for situational indiscernibility between the two land cover types (e.g. in mine or industrial construction sites). The final definition saw the inclusion of bare ground in construction sites and areas of compacted dirt, such as parking areas, being allowed in either the barren or impervious classes.

Challenges extended to other classes. In the shrubland land cover type, the initial definition of the shrub class was identified as an intermediate height value between the ranges of low vegetation and tree canopy land cover types. This value range was set between two and five meters. While seemingly logical, this definition did not account for the impact of old or sparse Lidar that did not match conditions in the imagery. This led a mismatch of what was visually confirmed to be shrub in the imagery compared to what the rulesets presented as shrub based on the definition provided. The definition was revised and updated to highlight the context-dependent nature of the shrub class, allowing for considerations in texture variability, height, spacing, and area covered (see Figure 1). Furthermore, this meant that time was spent going back to several sections that were classified under the original feature definitions and needed to be updated.

Figure 1. Example of updated shrubland definition.

Classifying structures in areas without Lidar presented another challenge. Structures and other impervious surfaces are spectrally similar; shadows also can be spectrally similar to these two features. Without Lidar, separating these becomes a tedious manual task. For West Virginia and New York counties without Lidar, extraction was sourced to Washington College's GIS team. Results of this exercise were not ideal. Segments generated from the NAIP and leaf-off imagery often were not precise enough to meet quality goals for the end product, in part because of the lack of Lidar, as well as because of the poor imagery quality (see Figure 2). It also cumulatively took the student team thousands of hours to complete the task.

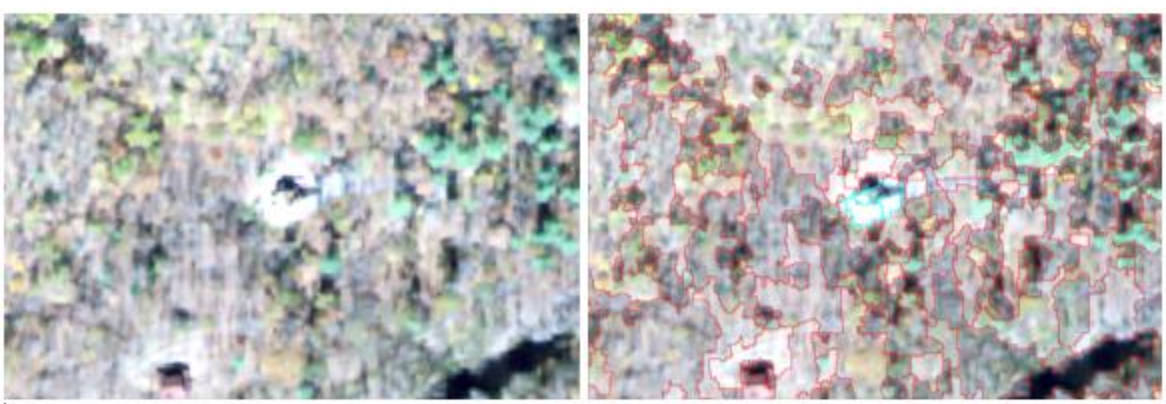

Figure 2. Segments highlighting a structure in West Virginia.

Logistics of processing and delivering the classification were prone to inefficiencies. Data was to be delivered on a rolling basis by county, with a 100-meter buffer of extra data to be processed around the edge of a county's boundary. This presented several challenges:

1) The rectangular quarter-quadrants of the NAIP tiles, which were the base processing unit of the classification, are not planned around county boundaries. For CC, which was using ENVI to perform its classification work, ENVI did not handle clipped imagery well, and analysts needed to segment and classify whole tiles, even if they intruded beyond the 100-meter buffer into another state where UVM or WVS was working. This created unnecessary redundancy classifying tiles in ENVI.

2) Planimetric data typically is planned on a county-by-county basis, but not with a 100-meter buffer around the boundary of the county, so data quality would differ within and outside the county boundary.

3) Lidar also often is collected and processed based on political boundaries. This dataset can vary county to county in quality and vintage, which makes merging of datasets difficult or impractical.

Another challenge was presented in combining products from the leaf-on and leaf-off imagery. Typically, these data do not align pixel-to-pixel, feature-to-feature, in collection timeframe, or in imagery collection angle. Merging results from these classifications led to a variety of issues such as overestimation of impervious extent and designations of tree canopy over the various impervious surfaces.

Finally, significant challenges were encountered as analysts compiled land cover products from the three contractors into one Bay-wide dataset. Data compilation involved bringing together the datasets across all the individual mapping areas along with those developed by each contractor. This process ended up being much more time-consuming than initially anticipated as gaps, discrepancies, and inconsistencies across these areas needed to be resolved.

Recommendations

Several steps could be taken to improve a potential project's workflow. In future initiatives, every effort should be made to set the mapping standards at the project initiation with pilot areas utilized to obtain CBPP feedback. The data delivery schedule should be shaped around expected component data updates. Additionally, land cover should be delivered in geographically contiguous blocks (which often do not align with data updates) or on a statewide basis, and should be checked continuously for seamlines between imagery tiles and political boundaries. A greater amount of time should be budgeted for correcting classes that are particularly important and that require viewing at a finer scale, such as impervious features. Where ancillary data is not available or does not meet quality standards, subcontracting those features to another party should be considered and planned for well ahead of time. Class definition and minimum mapping units should be established, critiqued, and revised prior to the start of a project. In the event that a feature definition is edited during the project, one person should be responsible for correcting the misclassifications in the sections that were completed under the original definition. This ensures that the project will continue to move forward in a timely manner, and that the misclassified features will be corrected in a consistent manner. Finally, a pilot study of hands-on classification time/error ratio should be done, as should an initial assessment of different classification softwares.

A future classification also should start with a comprehensive data management strategy. The volume of data that was processed was nontrivial, requiring high-performance computing and a considerable amount of online storage. It is likely that future land cover mapping endeavors in the region will have to ingest even larger data sets. Pre-project time should be spent planning for a data storage strategy, as well as on standardizing data preparation models. From the beginning of the project, all data should be projected to a single prearranged coordinate system and snapped to a uniform raster. If leaf-on and leaf-off imagery will be processed, the same person should be responsible for generating classifications from both. Furthermore, if planimetrics meet the quality standards, to save time, those features should not be classified with the ruleset. Instead, they should be wholly incorporated into the final product as part of the post-processing system. Another efficiency improvement strategy is to store segmentation files for QA/QC in a geodatabase, which, besides reducing draw time, also limits file size. In addition, versioned editing will help to avoid QA/QC vector data locks and file duplication. During QA/QC, standard definition queries should be utilized through ArcMap and ArcPro to find consistent errors across assigned geographies. This reduces the amount of effort spent panning and manually correcting errors.

Next, software and classification strategy could be improved. Ideally, a system for automated feature extraction for the Chesapeake Bay watershed would allow any geospatial data type to be ingested, permit batch processing, and yield high-quality output that is on par with manual feature extraction. The feasibility of this recommendation is uncertain within the timeframe (~2020) of a desired update to the 2013/2014 land cover data. Of the software packages used in this project, eCognition proved to be the most suitable, and given the multitude of problems encountered by CC in using ENVI, ENVI is not recommended. eCognition provides a range of more advanced processing capabilities that outweigh the greater initial upfront cost (in 2017, approx. $16,000 - $18,000 per permanent license). The purpose-built software is generally more adept at classifying imagery in regards to accuracy, aesthetics, and time. Another notable difference between these two programs is the lack of QA/QC interface in ENVI; eCognition contains built-in spatial and trait-based QA/QC abilities that cut down on classification time. Presently, ArcGIS's imagery classification software is being tested for its performance, and may present a reliable alternative to both platforms. Another improvement could be made to the classification methods by using a layered approach rather than a simultaneous one. In other words, CC initially classified all target classes in ENVI simultaneously, and implemented rule-fit thresholds to exclude classes during rule-based classification. Based on preliminary CC tests, and testimonial from UVM, a layer approach is more efficient. This would classify, for example, water and then remove those features from the classification, followed by tree canopy, and then impervious surfaces, etc., iteratively removing the categorized features as the process moves forward.

If ENVI (in 2017, approx. $2000 - $3,000 per permanent license) is selected for its upfront investment savings compared to eCognition, several recommendations should be followed. Vegetation indices such as the Normalized Difference Vegetation Index (NDVI) should be calculated through the classification software. Any datasets that are below a predetermined usability threshold should not be merged with the imagery to create the classification segments. For instance, Lidar should not be merged with the rest of the imagery that is going to be processed through ENVI if it falls below the threshold. ENVI produces temporary files each time it creates a preview of a classification, which rapidly consumes computer space. Any problems that stem from large intermediate files can be mitigated using programs such as WizTree to identify and discard those that are taking up the most disk space in one quick step. WizTree can be used only after ENVI has been closed, which means that an analysis machine needs adequate solid-state storage to accommodate this setting. Some other steps can be taken to reduce the amount of superfluous data produced through ENVI. Default outputs, such as creating floating point data for the millions of classified features, should be removed to increase user capacity of ENVI (see Figure 3).

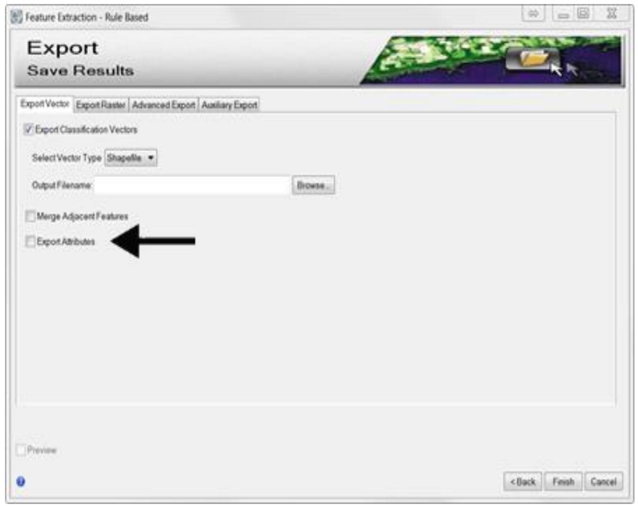

Figure 3. Optional attribute data.

Another significant challenge is that land cover errors propagate into subsequent land use classifications, which are based on the land cover datasets. Consequently, CBPP emphasizes the necessity of collecting and integrating imagery, digital surface models, and a variety of ancillary datasets to inform the land cover classification and to minimize misclassifications (e.g. classifying surface mines and rock outcrops being classified as “impervious”). To implement this, the land cover project team would undertake an explicit upfront data collection effort aimed at acquiring relevant data. Any future classifications also should consider whether it makes more sense to classify herbaceous into cropland, pasture, and turf as part of the land cover work or as part of a post-processing land use mapping exercise. The former may be preferable given the ability to leverage spectral data in the differentiation of herbaceous classes.

Source Data

Data used in the classification process.

In the interest of keeping to project deadlines, CC and UVM favored a mixed-source data strategy. This approach incorporated local, state, and federal vector data directly into the land cover in locations where they met quality standards. Depending on quality, data either wholly substituted for a class (e.g. structures) or was edited and then augmented remote sensing of that class. See this spreadsheet for a catalog of the source data used by CC.

Successes

By industry standards, the bulk of the classification work on this project transpired over a very short time for the amount and quality of data produced as well as the area covered – approximately 10 months. A mixed-source data strategy significantly cut down on time spent classifying and manually correcting features. NAIP proved to be a crucial data set for this project in that it provided a consistent spatial and temporal resolution for the mapping activities. Without NAIP, the final land cover product would have greater inconsistencies. Where available, Lidar improved quality of classification outputs and speed of corrections. It was crucial in the distinction of trees from herbaceous vegetation, and structures from other paved surfaces. Other derivatives such as Lidar intensity were valuable in distinguishing shrubs from trees, impervious from non-impervious, and water from non-water features. The incorporation of planimetric data into the classification process further improved results and facilitated smoother feature outlines in the final land cover classification. It also aided extraction and correction of wrongly classified land cover classes through feature queries and selections in ArcMap. Planimetric data such as driveways and sidewalks helped with detection of these land features in regions which were misclassified or not detected by ruleset schemes.

Challenges

Component data differed dramatically in quality and availability across the watershed. These data differed with respect to acquisition parameters, resolution, accuracy, and kindness. Preparation steps involved in such activities as generating image mosaics, normalizing and classifying Lidar point clouds, generating Lidar raster surface models, and reviewing vector data sets to confirm their quality. The most arduous task of the data preparation phase was compiling the data from all of the respective sources.

Collection years for Lidar often did not match with collection years for parent imagery (see Figure 4). Furthermore, the NAIP collection times within the 2013-2014 window varied state to state, with some having imagery collected in leaf-off, snowy conditions, rather than in the leaf-on period for foliage. State-collected orthophotos (i.e. high-resolution leaf-off imagery) often did not match the vintage of the NAIP imagery; origin dates ranged from 2005 in PA through 2015 in WV.

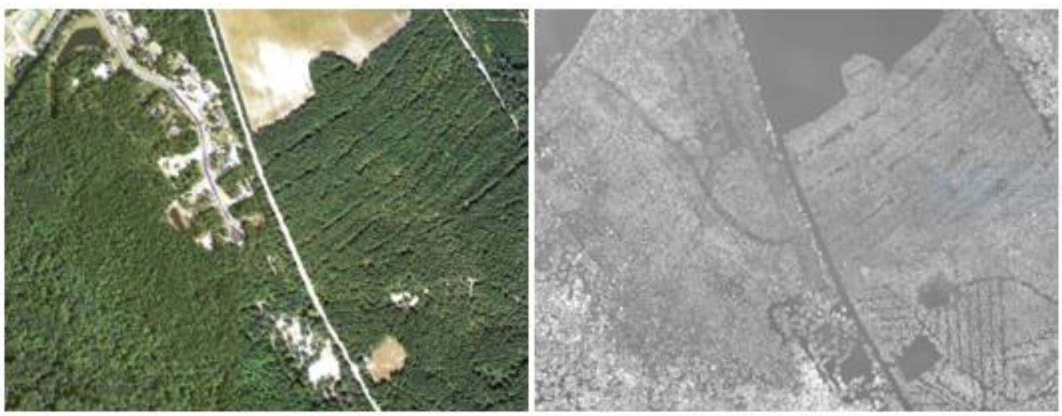

Figure 4. Difference in collection times between parent imagery (left) and Lidar (right).

Imagery and Lidar issues resulted in disparate class and geography accuracies, depending on the corresponding data limitation. For example, in West Virginia, the majority of the orthophotos used in the classification of impervious surfaces lacked the valuable near-infrared band. Additionally, pixels were missing from the imagery. The NAIP imagery for that state was collected late in the year; consequently, swaths of trees already were missing leaves. Furthermore, the color balancing made it difficult to discern barren from impervious surfaces, and due to flight angles, shadows often presented issues in identifying underlying features. Below are examples of the West Virginia NAIP imagery (Figure 5 and Figure 6), which was blurry and washed out. Figure 5 shows the scale at which manual corrections typically were done (right); at this scale it was hard to discern some of these features from others. Figure 6 is from the leaf-off imagery, and the red highlighting (right) indicates examples of missing pixels, which made it difficult to obtain usable segments for classification.

Figure 5. WV NAIP example (left) and at manual corrections scale (right).

Figure 6. WV leaf-off example (left) and highlighting missing pixels in red (right).

Other states had similar issues with imagery vintage and quality, especially in New York and Pennsylvania. Some of New York and almost all of West Virginia also suffered from a lack of Lidar to differentiate between tall and short features. Another problem occurred with the NAIP imagery, concerning the occasional lack of overlap for adjacent images. This created one- to two-pixel wide lines of NoData along tile seams, which had to be filled with time-consuming ArcGIS processes rather than a classification.

The quality of planimetric data presented another challenge. It varied widely across county and state boundaries, and the majority of the watershed was completely without the data, or lacked any high-quality information. Some of the challenges that detracted from overall quality of the planimetric data included but are not exclusive to the following:

- Omission of new developments, which is due to the lag time in updates to planimetric data (see Figure 7). Omission of new developments, such as residential and commercial structures, impaired the proper functionality of the aforementioned uses of the planimetric data. In this instance, classification and feature outline corrections using planimetric data proved less effective due to these data gaps and omissions.

Figure 7. Sections of road not included in planimetric data.

- Overrepresentation and underrepresentation of feature edges and outlines (see Figure 8). The overrepresentation of impervious features created exaggerations in the final classification, whereby features represented by the planimetric data encroached into classes adjacent to the planimetric data in the actual imagery. These exaggerations reduced the accuracy of the land cover dataset. A prime example of this occurred on road edges where the road or driveway layer datasets extended outside of the edge of roads and into the surrounding landscape.

Figure 8. Driveways are overrepresented in planimetric data.

- Spatial accuracy of planimetrics. For example, in sections of NY, WV, and PA, NAVTEQ planimetric data was used to help distinguish impervious road features. While generally accurate in its surface area coverage, NAVTEQ data was not accurate in its locational representation of road features, oftentimes failing to follow the edge of the road. Similar issues extended to structure outlines (see Figure 9).

Figure 9. Spatial inaccuracies in planimetric data.

Many other issues arose with the parent and processing data, including storage. Lidar, imagery, local vector data, and intermediate processing layers tended to be large and difficult to manage and store as they accumulated through the duration of the project. Server space filled quickly, and extra storage was purchased as needed; this strategy created a disorganized scheme that was cumbersome to navigate and manipulate.

Recommendations

The quality and speed of classification could be improved significantly with uniform, high-quality, high-resolution imagery that has at least four bands (red, green, blue, and near-infrared), and Lidar or a photon-derived Digital Surface Model (DSM) that matches the collection timeframe of the imagery. Should more consistent data sets be available in the future, the final Bay-wide data compilation task also would be correspondingly more straightforward.

A cost and usability comparison for classification and change detection should be conducted prior to purchasing. It also would be extremely beneficial to have standards for planimetric data that control factors such as quality and attribute conventions, or even setting a usability threshold for temporally mismatched parent data (e.g. Lidar and NAIP) that would provide a reference point for quality standards. If possible, a specialized team should take responsibility for obtaining and preparing all raw data in order to ensure a reliably consistent product to classify.

CC has been provided with price estimates for imagery and Lidar acquisition from Hexagon Geospatial (Hexagon); these may be subject to change. Contracting through Hexagon, the approximate cost of a coordinated data collection for the Chesapeake Bay watershed and surrounding counties is $20,200,000. Quality-level 2 Lidar costs $180-$190 per square mile, and 4-band imagery at 30-centimeter resolution is $12-22 per square mile extra, for a total of around $202 per square mile. An alternative to Lidar data is a Digital Surface Model (DSM) that is calculated with digital photogrammetry. These are typically much less expensive per unit area than Lidar data, and would be adequate for image processing support. Hexagon is able to create this type of DSM. However, the availability of the data is constrained by the purchasing history and plans of an anonymous "global client" that buys bulk orders of imagery. A Hexagon Representative as well as CC partners posit that the availability and price point of 30-centimeter resolution imagery is dependent on Hexagon's global client's interest in acquiring data at that resolution as part of the larger US NAIP acquisition program. For a county in Arizona that is partnering with CC on a land cover project, within the global client's acquisition area, staff have been quoted:

1. $10/sq mile 30cm 4-band imagery and $2.50/sq mile for 0.8M Nominal Point Spacing (NPS) DSM*

2. $20/sq mile 15cm 4-band imagery and $5.00/sq mile for 0.4M NPS DSM*

   * 10% price reduction from these levels for older imagery (e.g. 2015 products)

For a larger statewide AZ acquisition, the United States Forest Service and Bureau of Land Management have paid $0.50/sq mile for a 1.6M NPS DSM. Imagery price was not yet available. Other options to be explored include DigitalGlobe and Planet Labs imagery. In particular, future technologies such as single-photon Lidar may facilitate the availability of high-resolution 3-D data for the entire watershed.

Ideally, data acquisition for the Chesapeake Bay Watershed would be coordinated across all jurisdictions, but this scenario is unlikely to ever occur given the administrative and political diversity throughout the Watershed. If a coordinated collection of imagery, Lidar, and/or planimetric data is not possible, time should be invested in planning around the sources of component data, any pending updates to the data, and a products delivery schedule should be based on the collection timing. For example, if counties are intending to complete updates to planimetric data during the project period, completion of that county should be postponed to match the schedule of the update. Additionally, all component data should be assembled as far ahead of the project start date as possible. Other strategies for improving data include the development of a data storage plan, so that it can be organized in an efficient manner (i.e. planning hard drives for specific data types). Data also should follow a template that organizes temporary files from final files, preferably by state, and should be listed in one central document or spreadsheet that contains staffs' section assignments. Along these lines, a naming convention should be established with the data storage plan; this would further improve organization and ease of access by analysts. Additionally, instead of one uniform accuracy or data development time standard, separate standards should be applied to the land cover classes and to states depending on data quality and availability.

Data review and county requirement gathering

Interactions with county stakeholders to plan and review data deliverables.

Depending on local capacity, counties and/or state GIS staff were contacted for their input about the land cover classification; most conversations were done one-on-one via email or over phone. Comments for CC were submitted through customized ArcGIS Web Apps. Download CC's commenting methodology that was distributed to counties here, as well as for a summary of comments provided to CC here.

Successes

Engagement with county GIS and planning offices emerged as one of the most significant benefits of this project. The process sparked relationships with GIS users across the watershed, and created a network of interested parties at the local, state, and federal levels, as well as at non-profit organizations and universities. Each piece of the process contributed value. Curation of local vector data was useful for establishing contacts, and in the classification process for improving and expediting products. It also increased county office comfort with the final land cover, and paved the way for CBPP to reach out again with review requests for the land use data.

The attainment of county data ahead of time was also crucial to the success of the classification process because it permitted the analysts to appraise each dataset and choose how to move forward with their assignments. For instance, knowing ahead of time if planimetric data for a specific county corresponded well with the parent imagery, or not, helped the analyst plan how much of their time should be designated to each of their classification sections.

There was also a brief period of time allotted to involving the counties with the review process, which was beneficial in maintaining their engagement in the project. The series of web viewers that CC created for the counties to review the data was well-received. Upon final delivery, users overall were pleased with the data and appreciative of its capabilities.

Challenges

For all of its benefits, county engagement also presented roadblocks to the classification process. The most significant of these was the lack of established expectations about the land cover appearance and component data going into land cover review, as well as about the comment process more generally. Another was finding a contact person. Turnover at county GIS and planning offices contributed difficulties to establishing contact and a reliable reviewer, and also in obtaining planimetric data. Contacts were emailed between one and three weeks before data was distributed for review, which lasted three weeks. The webinar about county review occurred in early 2016; this was later than the first round of data review. CC also did not anticipate the burden that data review would present for county offices. Consequently, several weeks to one month was not enough lead time and three weeks was not enough review time for many, particularly for land cover drafts distributed around the Christmas holiday season. Once counties received the data, many were concerned about the pixelated appearance of the land cover classification, as well as its file size and their ability to manipulate and perform analyses on raster formats. In counties where local data was not found by CC/UVM, or by CBPP in advance of the project, there was displeasure at not having them incorporated. Furthermore, if planimetrics were available, counties often insisted that they be assimilated into the classification regardless of quality. Finally, this process was quite time consuming on the CC end, more than anticipated, and presented staffing challenges.

Recommendations

Several achievable steps can be taken to improve this process for future land cover classification iterations. More widespread county engagement about the whole project, both in advance of initiation and during data creation, is needed, so as to reduce gaps in communication. Project leaders should share with counties or other stakeholders, well in advance, any expectations for involvement, as well as land cover samples, and a realistic review schedule. Another strategy could be to designate one dedicated staff person to manage all county outreach and engagement through the duration of the project. This would ensure that adequate focus and attention is paid to the process and avoid bottlenecks with otherwise-engaged project staff.

Accuracy and deliverable expectations

Accuracy assessment development and client response to product accuracy and general rollout.
An innovative protocol was devised to tackle accuracy assessment issues that arose when applying traditional methodologies to the high-resolution land cover classification. This new process ensured a fair assessment that could be translated to future iterations of high-resolution classification projects, and was thoroughly tested to prove its transferrable nature.

The results of the accuracy assessment averaged 90% for the entire watershed when measured on a statewide basis. This was met with CBPP approval given the context of unforeseen complications throughout the project, such as fluctuating feature definitions, which compelled CBPP to lower their initial class-by-class accuracy expectations. CC's public website hosts the final land cover classification and accuracy assessment methodology for any stakeholder to access the information. Download the full methodology here.

Successes

The workflow of CC's final accuracy assessment protocol boasted several successes. One of these was leveraging ArcGIS Pro, which allowed analysts to link the views (i.e. extent and scale) of two maps, one each with leaf-on and leaf-off imagery. Thus, reviewers were able to consider both sets of images when determining the correctness of a sample, rather than toggling between images within a single map. This saved valuable time and sped the process toward data sharing. Additionally, importing publicly-available imagery services for parent imagery saved time over making mosaic datasets of the thousands of image tiles. Another success of the process was incorporating the input of at least two analysts for each accuracy assessment sample to establish agreement about the actual land cover class. This boosted confidence in the assessment results, provided a detailed record of incorrect data, and increased understanding of areas of confusion between classes. Finally, users of the CC land cover project website have commented that the interface and openness of the information is valuable.

Challenges

Accuracy statistics varied dramatically with assessment methodologies, and CC tested several iterations of the accuracy evaluation protocol before determining a final version. This created some uncertainty during the project about the accuracy of the land cover data, classes that needed correction versus those that were meeting standards, and consumed valuable time in creating the trial assessments. Furthermore, it was difficult to reconcile several of the aforementioned challenges in the accuracy measurement:

1. Variable parent data selected for use in the classification. The incorporation of Lidar, leaf-off imagery, and vector data varied county-to-county depending on quality and availability. CC decided to evaluate only imagery, and to establish consistency between the NAIP and leaf-off imagery before evaluating an assessment sample.

2. Misalignment of leaf-off imagery and NAIP imagery. As explained above, leaf-off imagery pixels did not align spatially or temporally with NAIP imagery. This impacted the observed ground conditions of the assessment samples. Again, CC established consistency between the NAIP and leaf-off imagery before evaluating an assessment sample.

3. Changing land cover class definitions. CC persisted in evaluating the accuracy against most current definitions, but permitted some confusion between classes when definitions were particularly inconsistent across states, such as with shrubland and tree canopy, and barren

land and impervious surfaces.

Difficulties also arose in navigating the organization of deliverables to multiple layers of clients: both to CBPP and the counties. The rolling deliverable timeline was not ideal because it was very complicated to manage. Any successive corrections required that the data be resent to the clients, which created opportunities for miscommunication. Counties also were not as familiar with the raster data format, which prompted many questions as to the pixelated appearance of the final product. Additionally, CBPP did not invest in soliciting input from counties in the project development process, creating confusion regarding the project's context when trying to incorporate county feedback. Questions concerning CC/UVM's mission, project timeframe, and technical aspects arose from this issue. While CC and UVM met the expectations established by CBP, they remained accountable to the counties as well, and had to navigate through these types of inquiries.

Recommendations

Expectations for accuracy should be based on the quality of imagery and component data, as well as the project timeframe. A protocol needs to be agreed upon with CBPP before the project start that defines the number of samples to be gathered and the appropriate area of accuracy analysis, preferably by state, and any other relevant methods considerations. A fuzzy accuracy assessment is the most ideal methodology for this type of data, and should be conducted impartially without looking at the classification. Samples are to be based on the different land cover classes and should not intersect with transition zones from one class to another.

Future land cover classification projects should try to avoid rolling deliverable timelines, and rather plan to send all data drafts out to clients at the same time. Quality expectations with all involved parties should be set well in advance of the project start date and deliverable deadlines, and clearly communicated with all clients. Additionally, CBPP should be involved in initial contact with counties to further facilitate communication regarding project details. This includes coordinating ahead of time in regards to contacting counties and assisting with introductions, making sure everyone is on the same page, and explaining the purpose of the project, how it will benefit clients, and deliverable appearance expectations. Better publicity about the project context and its legitimacy as an official measurement dataset is imperative.

Part 2: Feedback from Stakeholders about the Chesapeake Bay Watershed 2013/2014 High-resolution Land Cover Classification Project

Introduction

CC has spearheaded an outreach campaign to amass existing and emerging use cases, gauge usability, and collect feedback about valuable resources and edits for any updates to come.

Specifically, deliverables that have been requested by CBPP are as follows:

- Most commonly used land cover classes
- Feedback from government agencies and organizations about existing/emerging use cases
- Identification of missing or inadequate datasets relevant to land cover

Additionally, CC has identified recommendations for the following:

- Obstacles to applications of land cover data
- Recommendations for outreach and engagement moving forward
- Complementary layers, datasets, and tools
- Useful resources/tools (webinars, web applications etc.)

Deliverables have been collected through a combination of surveys and interviews. The following section will go over outreach and engagement methodology that has been implemented, a brief profile of the respondents, and provides links to appropriate resources, such as the surveys disseminated.

Outreach and Engagement

CC's approach to identifying the myriad of variables listed above began by cultivating a list of appropriate and targeted professionals to contact. Thus, the campaign was started by creating a matrix of professionals involved in conservation and restoration work throughout the watershed, ensuring that each watershed state had representation from the federal, state, and local levels, as well as from the non-profit and academic industries. Most contacts were taken from the contact list provided to CC by CBPP at the start of the Land Cover Data Project. Other contacts were taken from High-Resolution Land Cover Webinar respondents, and from existing partnerships.

From here began the email campaign. Approximately 300 people were reached through targeted emails that were tailored based on factors such as the recipient's nature of work or past partnership with CC or CBPP. For example, during other outreach efforts people were polled on their level of interest in receiving news or providing feedback to CC about the high-resolution land cover dataset. These respondents were batch emailed together using Constant Contact so as to: 1) manage organized lists of contacts efficiently, 2) protect the contact information of those involved, 3) allow recipients the ability to unsubscribe from future outreach campaigns, and 4) provide basic analytics about CBPP's audience. Other groups that were batched together included staff involved in CBPP's Land Use Work Group, participants in the West Virginia Eastern Panhandle GIS Conference where CC presented, GIS users involved in the Chesapeake Conservation Partnership, and contacts referred by internal and external colleagues, to name a few.

These targeted lists received one to three emails encouraging them to participate in either a survey, interview, or other sort of informal exchange in order to learn more about the nature of their work. Those that responded were followed up with individually, asked to fill out a general questionnaire, and then interacted with based on their preference of communication and contact (i.e. phone, email, Google form etc.).

Of the 300+ people who were contacted, 36 responded to the surveys, and approximately 26 individuals from varying fields of study and geographic location were interviewed. General demographics, job title, and industry of those who responded can be viewed in the **Survey Figures** section (Figures 1-7) of **Appendix A**.

State representation for respondents was lacking in Washington, D.C. and New York for both interviews and survey respondents, and representation could have been more robust in West Virginia for interviews. Despite this, a diverse and fairly representative set of participants was engaged across the watershed. Additionally, almost equal representation was secured between the state, local, non-profit, and academic arenas. Engagement likely could be improved with more targeted messaging from both CBPP and CC, and a more vigorous effort to coordinate messaging and channels with respective programs and communication teams.

To see an example of questions respondents were asked either over the phone, or via a Google form, see **Appendix A**, or the sample form attached here.

Existing and Emerging Use Cases

All individuals (i.e. survey respondents and verbal interviewees) were asked if and how they use land cover data in their work, or how it was relevant to their work. Major themes included: habitat analyses; riparian buffer and stream restoration site selection, evaluation, and prioritization; plans for stormwater work; and nutrient and sediment load calculations. The examples below illustrate how a couple of respondents apply land cover data in their work:

"We use land cover data to examine the watershed draining to our stream restoration sites. This allows us to model the hydrology and understand possible causes of impairment. We often use aerial photos combined with StreamStats to estimate watershed conditions, but true land cover data is more valuable, and I use it to check our consultants' estimates."

"We use land cover data to help design and identify potential stream restoration projects. Land cover helps assess stream health, and estimate stream discharge during different rainfall events."

For a full list of respondent answers, please see **Table 1** in **Appendix B**.

During the interviews, respondents were asked to elaborate about their use of land cover data. In particular, they were asked if and how they were using CBPP's high-resolution land cover dataset. Most had not yet had a chance to incorporate CBPP's dataset into their work given how recently it was made available. However, several existing and emerging use cases were discovered. These included, Sarah Johnson, from the Harrisburg Chapter of The Nature Conservancy, PA; Cory Sauve, a student working on a thesis project for Mansfield University in Pennsylvania & Scott Zubek, the Director of Tioga County, PA GIS; Steve Miller, working in Adams County, PA; Adrienne Gemberling, Susquehanna Technical Coordinator, and finally, Alex Metcalf & Conor Phelan, University of Montana.

Case Study 1: Translating CBPP Land Cover Data for CBPP's nutrient and sediment load modeling across watersheds

Sarah Johnson, Conservation GIS Analyst, The Nature Conservancy

Sarah consistently works with land cover data, separating the nature of her work into two main categories: prioritization modeling and scoping for future conservation opportunities, specifically in regards to intact forest and agricultural lands. She relies heavily on NLCD for land cover data, especially when working on projects at the landscape scale, and occasionally on PA LULC, and most recently, utilizing the new one-meter high-resolution land cover dataset for the Chesapeake Bay. With the caveat that the nature of her work is highly dependent on each project, tree canopy/forest, wetland, and road classes were most important and common to her everyday work.

Specifically, she used CBPP's high-resolution land cover dataset while working on a project that aimed to measure nutrient and sediment contributions using the CBPP 5 nutrient and sediment load model. Out of curiosity, she did the analysis twice: once using NLCD (30-meter resolution land cover data), and once using the Chesapeake Bay Watershed land cover data (one-meter resolution) to see if there was any significant difference. Surprisingly, she discovered little to no major differences. However, she mentioned this may be because the model she is familiar with and uses in her work is calibrated to NLCD classes, thus before she could compare the differences she had to translate the high-resolution land cover dataset into like classes. This undertaking was not found to be particularly intuitive or quickly translated, so she informally translated both the one-meter land cover classes, and the CBPP land use categories. Given the difficulty of translating the data sets, she concluded that the result was unsatisfactory because it required leaving out land use information and classes, and "watering down" the precision of the one-meter resolution land cover dataset by combining it with the CBPP LULC dataset.

Figure 8 is an example of how she translated some of the classes. Land cover spatial classes are on the left and are color coded with the CBPP classes on the right.

	Uvm 1m land cover categories		Bay_Prog_abr	Bay_Prog_def	UVM_cat	NLCD_cat
Value	LandCover		hwm	high-till with manure		Cultivated Crops
1	Water		nhi	high-till with manure nutrient management		Cultivated Crops
2	Wetlands (emergent)	just for	hom	high-till without manure		Cultivated Crops
3	Tree Canopy	highest	nho	high-till without manure nutrient management		Cultivated Crops
4	Scrub-Shrub	average	lwm	low-till with manure		Cultivated Crops
5	Low Vegetation	highest	nlo	low-till with manure nutrient management		Cultivated Crops
6	Barren		hyw	hay with nutrients		Hay/Pasture
7	Structures	highest	nhy	hay with nutrients nutrient management		Hay/Pasture
8	Other Impervious Surfaces	average	alf	alfalfa		Hay/Pasture
9	Roads	highest	nal	alfalfa nutrient management		Hay/Pasture
10	Tree Canopy Over Structures	average	hyo	hay without nutrients		Hay/Pasture
11	Tree Canopy Over Other Impervious Surfaces	average	pas	pasture		Hay/Pasture
12	Tree Canopy Over Roads	average	npa	pasture nutrient management		Hay/Pasture
			trp	pasture corridor		Hay/Pasture
			afo	animal feeding operations		Cultivated Crops
			urs	nursery		Cultivated Crops
			puh	high-intensity pervious urban		
			pul	low-intensity pervious urban		Developed, Open Space
			imh	high-intensity impervious urban		
			iml	low-intensity impervious urban		
			bar	bare-construction	Barren	Barren Land
			ext	extractive		not used
			css	combined sewer system		not used
			for	forest	Tree Canopy & Wetlands	Deciduous Forest
			hvf	harvested forest	Tree Canopy	Deciduous Forest
			wat	water	Water	Open Water
			ps	wastewater		not used
			septic	septic		not used
			atdep	atmospheric deposition to non-tidal water		not used

Figure 8: Example of translated land cover classes.

The biggest piece of feedback received about this particular experiment was that although comparing the two different land cover datasets was not quite a success; she did not think it was necessary to mimic NLCD completely. Suggesting only specific classes be broken into sub-classes to match NLCD: wetlands, agricultural, and tree canopy.

She projected that an emerging use case for the high-resolution land cover dataset would utilize the road class to assist with a hydrologic function analysis—seeking out variables such as where floodplains are cut-off by development, or to help track aquatic organism movement and crossings.

Case Study 2: Utilizing CBPP Land Cover Data for Brook Trout Habitat at the Watershed Scale

Cory Sauve & Scott Zubek, Student, Mansfield University; Director of GIS, Tioga County

Cory Sauve is a recent graduate of Mansfield University who is involved in a project that strives to identify if there is a relationship between large-scale watershed factors such as tree canopy coverage over streams and unimpaired Brook Trout habitat in the Susquehanna River watershed. His study predominately utilizes the tree canopy, impervious surface, and roads land cover classes. Part of his study involves comparing his analysis with, "more common land cover resolutions (e.g. 30-meter resolution) to see if there is a significant difference in results."

At the start of this project, Cory reached out to the Tioga County, Pennsylvania GIS department to access data layers. Scott Zubek, the Director of GIS there, initially helped Cory find the necessary data layers he required and introduced him to the CBPP land cover dataset. From here, the two did a test study to analyze the difference in amount of habitat CBPP's dataset would "grab" versus University of Delaware's 10-meter resolution land cover dataset.

One of the factors Cory was interested in analyzing was shade coverage over streams, because Brook Trout prefer cooler water. His goal was to find a correlation between Brook Trout presence and stream canopy coverage as areas to prioritize for conservation and restoration projects. To do this, he started with an analysis that separated out tree canopy from the other classes, and then found where tree canopy overlapped with a stream layer.

Although his team found noticeable differences in the amount of tree canopy picked up between the two datasets—specifically noting how impressive the difference was between 10-meter and one-meter resolution data in urban and residential areas where classes tended to be more fractured—they noted that the size of the dataset was prohibitive to further analysis due to time constraints and processing power available to them.

Case Study 3: Envisioning the Battle of Gettysburg

Steve Miller, Executive, TerraConFirma Consultants

Steve is working on a historic forensic tool that will help users simulate how the Battle of Gettysburg unfolded. He most recently dedicated his time to creating a historic land use/land cover GIS overlay for the area of the battlefield using historic map surveys conducted in 1868-69, and has reprojected and georeferenced them to match modern orthophotography. He has relied heavily on National Agriculture Imagery Program (NAIP) to do the georeferencing and has utilized sources that represent the area in both leaf-on and leaf-off conditions. In addition to these sources, he is also relying on roads and cultural data, his own photos from 2004, and on-site visitation to accurately depict the landscape characteristics and features. He is now utilizing the high-resolution land cover data for Adams County, Pennsylvania to verify and ground-truth factors such as: areas that have remained forested, shapes and locations of buildings, and to check his own work--ensuring he has correctly interpreted the historic maps. An example of this is demonstrated in Figure 9 below, showing how his original fence overlay matches the natural fracturing across land cover data classes almost identically.

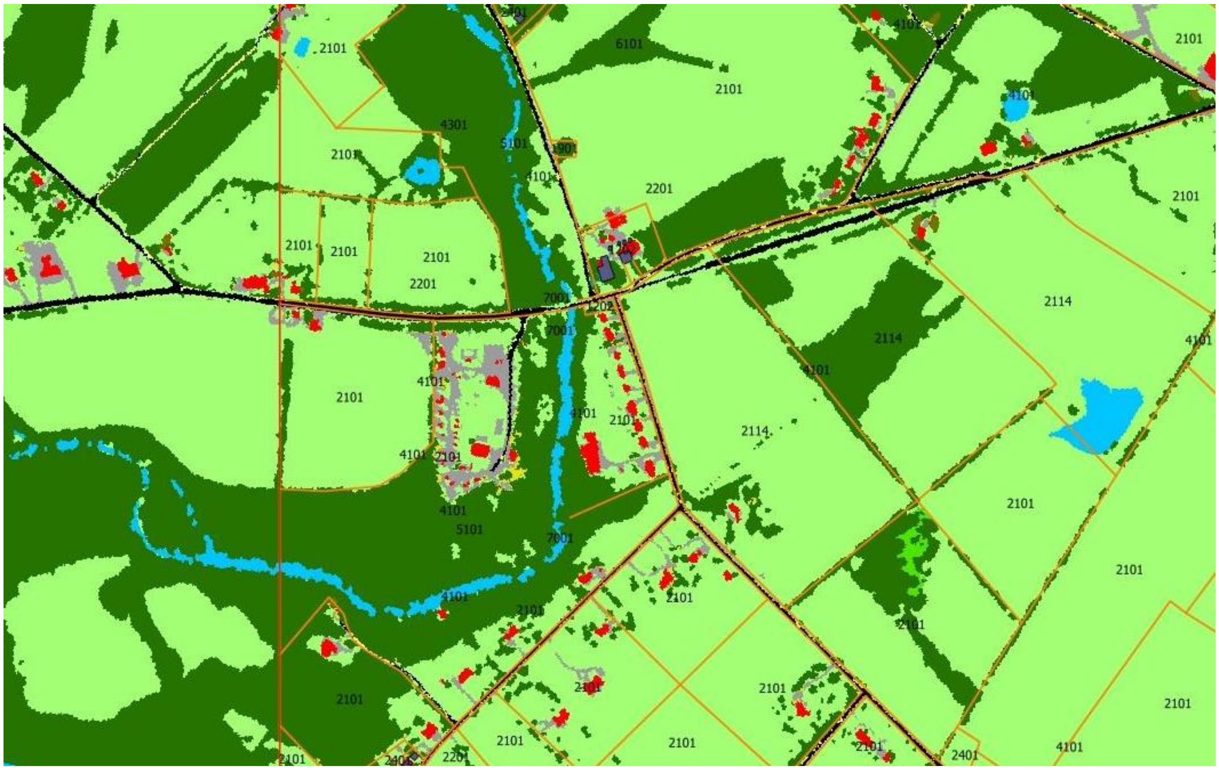

Figure 9: CBPP land cover in Adams County, Pennsylvania with fence overlay (orange).

Land Cover Categories Feedback

Out of these interviews and use cases it emerged that the most common and relevant land cover classes to respondents' work varied (as would be expected), but that tree canopy, wetland, and impervious surfaces were cumulatively the most commonly used classes. Figure 10 shows that, although each survey iteration and interview pool had different numbers of respondents, answers still followed similar trends, peaking at tree canopy, and again at impervious surface, and dipping along barren and structure classes.

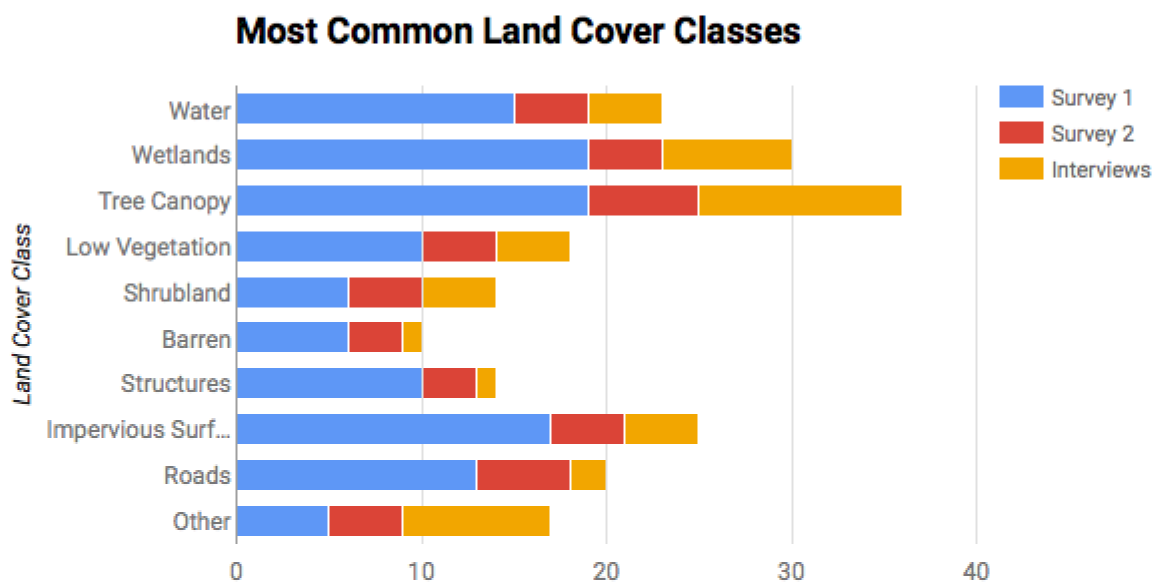

Figure 10: Most commonly used land cover classes by survey respondents.

Where respondents answered "other", it was asked that they specify, with the majority of respondents writing in variations of agricultural classes such as row crops, pasture, cover crops, and farmsteads. Other unique responses included tile drains and vernal pools. At least five of the 17 responses for "other" simply voiced their support and excitement for the tree canopy over impervious, structures, and roads classes. In an interview with representatives from The Nature Conservancy, their staff saw immense potential for these types of classes being used in tree connectivity analyses and for enhancing roads layers which they mentioned often was lacking compared to CBPP's land cover dataset.

Additional classes and complementary layers that respondents identified as "inadequate or missing" consistently fell upon the existing wetland and shrubland classes, with many voicing a need for a more robust breakout of these classes and mentioning a lack of quality resources across the field for identifying and mapping various kinds of wetlands and shrubland areas. One hundred percent of respondents indicated that the lack of an agricultural class fell into the "inadequate or missing" category.

Obstacles and Limitations

Some respondents have also provided feedback that they only "sometimes" use land cover data (of any kind) in their work, and others responded that they had not yet had a chance to utilize the CBPP dataset specifically. Indeed, our limited sample size is largely due to the narrow amount of time that has passed since the dataset was initially released and when people were contacted. Due to this, respondents were asked more general questions about land cover data in addition to questions specific to CBPP's land cover dataset.

When asked why they may currently not be working with, or using, land cover data in general including CBPP's dataset, participants responded with some of the following:

"Our work is very project focused, so unless I get a project which focuses on or would benefit from this information I don't generally use it...Unfortunately only half our service area is in the Chesapeake Bay watershed as well, so if despite the universality of land cover data, your dataset only covers part of regions I need it to."

"Staff time to develop GIS-based estimates relevant to our work, capacity to do updates and change analysis for trend tracking, limitations of hydrology layer, lack of information about groundwater inputs to streams, processing capacity internally for fine-scale data at statewide scale."

For a complete table of direct answers from participants, see **Appendix B, Table 2.**

When asked about frustrations specifically regarding high-resolution datasets, many respondents said that they initiated analyses understanding time and processing power would be challenges; however, time and computing constraints still hampered progress and ability to complete certain projects. Other examples of answers received include, but are not limited to:

- Slow loading
- Overcoming the "chicken pox effect" and continued shadow problems with classification of trees and forests*
- The resolution of the classes and lack of forest and wetland delineations is disappointing
- It seldom extends into WV, regional products are wonderful for the surrounding states, but we often find that data products do not extend into WV.

For a full list of direct responses please see **Appendix B, Table 3**.

*The respondent who used the phrase "chicken pox effect" was referring to the pixilated appearance of the dataset given its resolution and the "freckled" appearance that plagues forests and areas where shadows are misclassified.

In sum, respondents indicated a range of limitations to using high-resolution data which can be categorized into two different categories: an obstacle that arises out of a general capped resource, and an obstacle that was encountered specifically with the CBPP dataset.

General restrictions on resources cited:

- Quality of available data in any sense to be used for complementary or supplementary purposes. An example of this that was often cited was quality of agricultural data available through USDA.
- Time, many people said they could not afford the time to look for new high-resolution datasets, familiarize themselves with it, and use it for models or analysis due to processing time.
- Capacity in general (in terms of staff time, technical resources, training etc.,)
- Financial resources and budgetary constraints. An example of this was a GIS professional who performed all GIS support for every city department. Although time was cited as the biggest obstacle, a lack of training and proper equipment was named as detracting from the ability to reap the benefits of a dataset such as CBPP's.
- Slow loading and processing (both in terms of time and equipment available)
- Consistency—datasets being used together are of varying resolutions, times of year, sources, etc.

Restrictions specifically mentioned by respondents in regards to the Chesapeake Bay Land Cover Dataset:*

- Slow loading and processing time
- "Chicken pox effect," or issues with shadows and classification of tree canopy
- Dataset size and storage capacity
- Only partial coverage for states of NY and WV. All respondents from WV and NY State indicated that only having partial coverage of their respective states was prohibitory to their work.
- Class accuracy. Many users indicated needing more accurate and specific classes for wetland and shrubland classes.
- Consistency—some respondents were wary of the many different source years of imagery used and thought this type of dataset would only truly be beneficial if updated at consistent intervals.

  *Examples such as these are not to be interpreted as issues that other high-resolution datasets are immune to, simply repeated obstacles cited by respondents when speaking directly about the CBPP dataset.

The last major concern that was consistently brought to attention over the course of this study was understanding the accuracy of this dataset, as well as the year that it represents given the mixed-data classification approach.

In addition to being asked about obstacles to their work, and frustrations specific to the CBPP dataset, respondents were asked to answer the following questions:

1. Are there analyses you cannot currently perform? If so, why?
2. How can we improve our [CBPP's] land cover dataset?

Tables 4 and 5 provide sample responses to each question, and can be found in Appendix B. About 50% of the answers to question number one are also related to answers provided earlier about obstacles and frustrations: time, quality, and access to adequate resources. Other answers indicated a lack of need to incorporate the data given the nature of projects or requests coming in from stakeholders, clients, and the community. Answers to question number two, reiterated thematic concerns about class consistency across state lines, data extent not encompassing the full state of New York and West Virginia, and lack of information about agricultural fields. Others simply responded that they did not know enough about the dataset to comment, or skipped the question all together. Interestingly, 100% of people we interviewed requested that agricultural information be included in an update to CBPP's land cover dataset; and no respondents had been aware of the 10-meter resolution CBPP Land Use dataset that did incorporate agricultural information.

Recommendations

Based on this feedback, recommendations are as follows regarding dataset alterations to meet users' needs:

1. Include the land use classifications for agricultural land in the land cover dataset
2. Separate wetlands class into subgroups that mimic NLCD's i.e. woody wetland, emergent wetland etc.
3. Include an outcrop and mine class or layer, particularly for the state of WV
4. Separate tree canopy class into subgroups that mimic NLCD's i.e. deciduous forest, evergreen forest, and mixed forest
5. Create a more robust shrubland class
6. Expand NY and WV to state boundaries for state-wide analysis
7. Convert from raster to vector format for users
8. Consider a coarser-resolution dataset (e.g. 10-meter cell size) to help balance processing time
9. Calculate land cover class areas and percentages in the attribute tables.

Many respondents were also interested in additional resources, such as an updated web application, similar to the one hosted on CC's website currently, which would include options like checkable complementary layers such as: topography, elevation, slope, and general soils data. Others desired forums through which people could find common resources, such as a list of all the websites users could download the CBPP land cover data.

For example, many interviewees were confused about the details and hosting locations of the CBPP land cover dataset and CBPP land use dataset. Confusion and need for clarification between the CBPP land cover and land use datasets were pervasive issues throughout the outreach process.

Moreover, those who were aware of the differences and had listened to the webinar regarding methodology indicated that a similar webinar focused solely on example uses of land use data and how to manipulate the CBPP dataset would be particularly helpful.

Lastly, due to the high volume of comments regarding resolution and consequential size, and difficulty comparing the CBPP land cover dataset to NLCD land cover classes for modeling purposes, the following two questions were asked:

> Q: Would you be more satisfied with a 10-meter resolution vs. one-meter resolution given the balance between precision, size, and resolution?
>
>> A: Overwhelmingly answered, "yes" across positions and industries. For example, many land trust and county GIS professionals interviewed had no need for detail at that level, and found the high-resolution to appear pixilated, messy, and "unfinished". The GIS professionals who do more sophisticated analyses and modeling cited reasons regarding time constraints and processing power. The solution many saw as a compromise was to

calculate percentage and area of land cover class at the one-meter resolution, but package the publicly available dataset at nine- or 10-meter resolution.

Q: What classes from NLCD would you like to see mimicked or transferred into this dataset?

A: Many respondents were excited about a high-resolution land cover dataset, particularly in comparison to the coarse, but common, NLCD. An obstacle that arose when attempting to substitute the CBPP land cover data was that most models have been configured and developed based off NLCD classes, which do not match. Thus, one-hundred percent of respondents said there was no need to mimic every single land cover class included in NLCD, but identified the wetland groups, forest groups, and agricultural classes as being high priority in their work and ideal classes to mimic in updates to come.

Resources

To address some of the recommendations above, respondents were also asked to list channels, networks, datasets, and tools that were both relevant and common in their work. In the future, these resources could be integrated into the delivery of CBPP data and products.

Commonly used channels and networks included:

- ESRI newsletters, ESRI.com etc.
- WV association of Geospatial Professionals (WVAGP@listserv.wvu.edu)
- Chesapeake Bay Conservation groups
- Data hosting sites such as: CUGIR, NY GIS Clearinghouse, iMAP, WV GIS Clearinghouse, VIMs
- Chesapeake Network

Layers and datasets respondents specifically mentioned as being complementary and/or supplementary to their land cover related work included, but was not limited to, the following:

Table 6: Commonly used datasets in land cover work.

National Wetland Inventory (NWI)	National Land Cover Database (NLCD)	National Hydrography Dataset (NHD)
FEMA Flood Zones	Soils Survey Geographic Database (SSURGO)	Trout Waters
Elevation (DEMs)	Hydrology in general (surface water, depth of water table, water quality, watersheds)	Slope
Cadastral	Jurisdiction & Tax Parcels	Ecologically Significant Areas
Easements & Reservations Data	Historical Aerial Imagery	Impaired Streams Data

Similarly, the following websites were identified by respondents as valuable and/or frequently visited for information, as well as resources for data downloads:

Table 7: Commonly used data and news sources.

Name	Website
USDA Geospatial Data Gateway	https://datagateway.nrcs.usda.gov
Natural Earth Data	http://www.naturalearthdata.com
The National Map	https://nationalmap.gov
The U.S. Government's Open Data Portal	https://www.data.gov
Maryland's Mapping and GIS Data Portal	http://imap.maryland.gov/Pages/default.aspx
NatureServe	http://www.natureserve.org
STORET (STOrage and RETrieval Dashboard)	https://www.epa.gov/waterdata/storage-and-retrieval-and-water-quality-exchange
Maryland Biological Stream Survey (MBSS) StreamHealth Map, MD DNR	http://dnr.maryland.gov/streams/Pages/mbss.aspx
North Carolina State University's Stream Listserv, Water Resources Research Institute	https://wrri.ncsu.edu/contact-us/listservs/
NYS GIS Clearinghouse	https://gis.ny.gov
CUGIR	http://cugir.mannlib.cornell.edu
WV GIS Technical Center	http://wvgis.wvu.edu/data/data.php
County Departments for Planning & Zoning	Specific to county
SSURGO, NRCS	https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_053627

Conclusion and Future Steps

Although this specific outreach endeavor has been rich in information, it is safe to conclude that the high-resolution land cover dataset simply has not been publicly available long enough to obtain a complete picture of how stakeholders throughout the watershed are, or will be, utilizing it, and how they hope to use it. However, there has been no lack of enthusiasm for the opportunity to utilize the dataset. By increasing outreach and engagement efforts, and making a few adjustments to land cover classes in future updates, more organizations and partners would benefit from the land use and land cover datasets.

Respondents overwhelmingly said that outreach and engagement prior to the rollout of this dataset were lacking and weak points in successfully making the land cover a more widely utilized tool. Thus, if there were one major lesson to take away from this report, outreach and engagement efforts should be more robust, thorough, thoughtful, and coordinated moving forward. For recommendations regarding successful outreach efforts see **Appendix C.**

Appendix A
Survey Figures 1-7

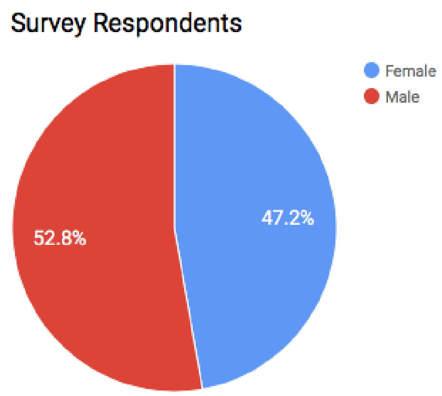

Figure 1: Gender of survey respondents.

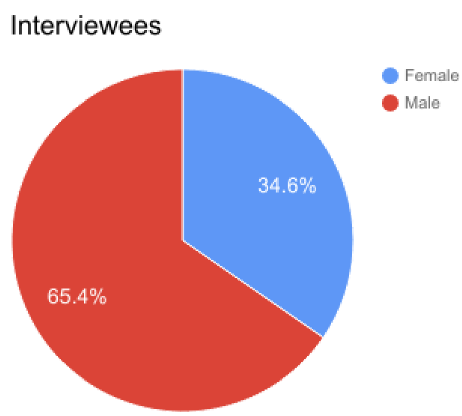

Figure 2: Gender of interviewees.

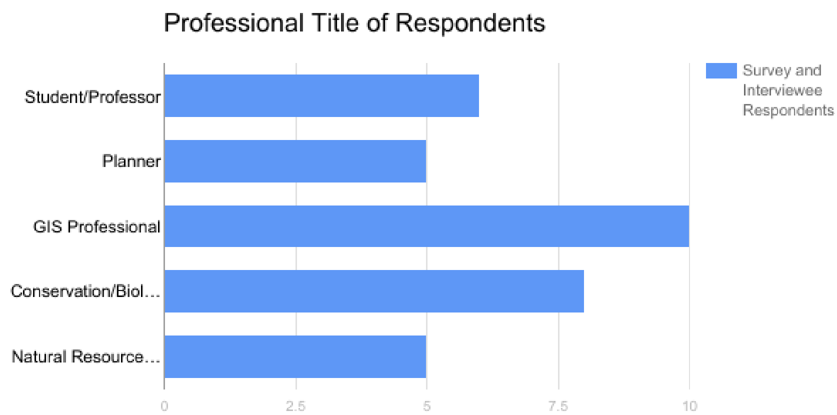

Figure 3: Professional title of all respondents (i.e. interviewees and survey respondents).

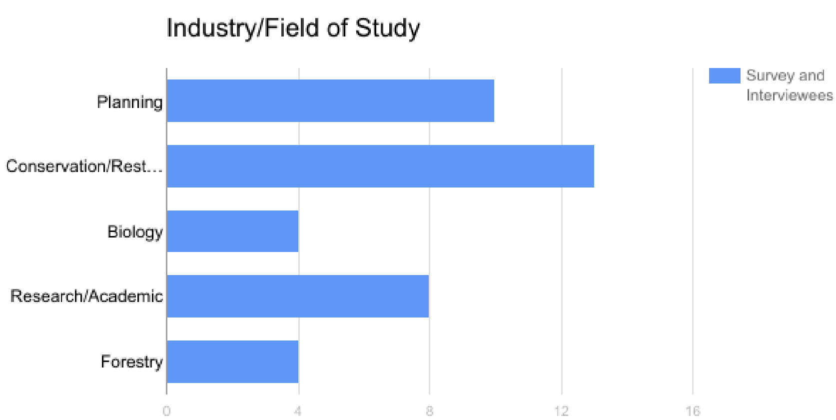

Figure 4: Industry/field of study of all respondents (i.e. interviewees and survey respondents).

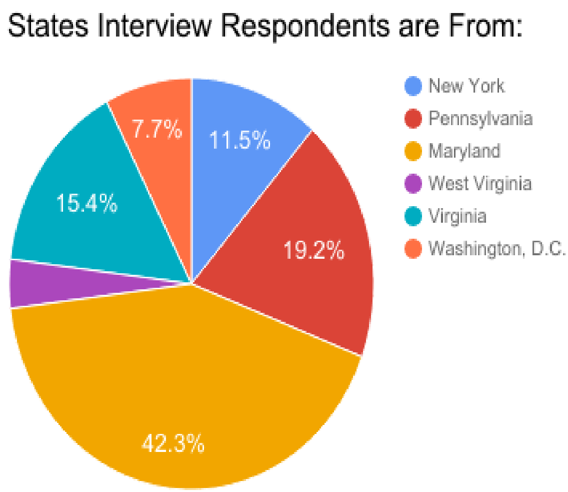

Figure 5: Breakdown by state of employment or residence of interview respondents.

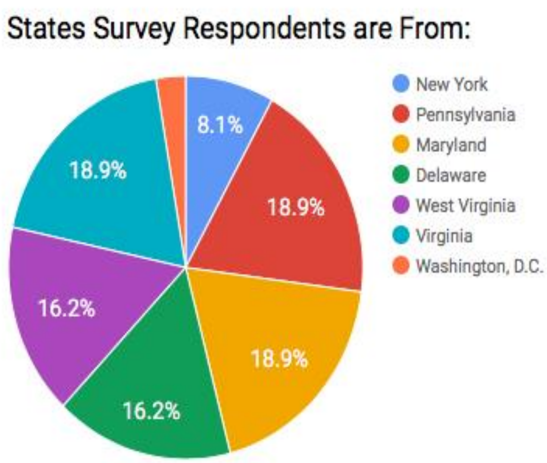

Figure 6: Breakdown by state of employment or residence of survey respondents.

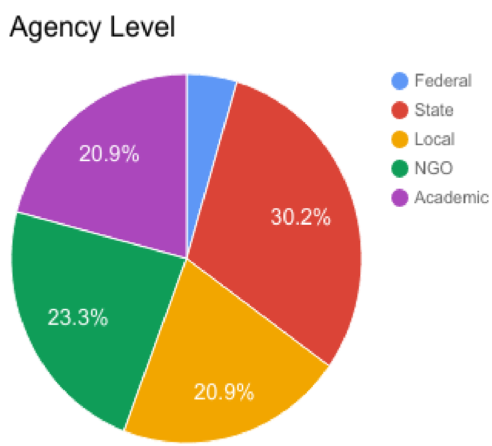

Figure 7: Government agency or other professional affiliation of interview respondents.

Survey Questions

Questions all respondents were asked either over the phone, or via a Google form, after agreeing to initial outreach are listed below, with an asterisk demarcating a required field. A sample is attached here.

1. Name*
2. Title*
3. Are you affiliated with federal govt, state level govt, local govt, academic institution, NGO, for-profit, other?
4. Please list your industry/department
5. Do you use land cover data in your work? Y/N*
   a. If yes, please explain the project*
   b. If no, do you have future plans to?
7. Please explain either past uses or future plans for land cover data
8. What are obstacles to your work?*
9. What are common GIS analyses you perform?
10. What are GIS analyses you would like to perform?
11. If there are analyses you cannot perform, or do not currently perform, why not?
12. Have you watched our land cover webinar?*
13. How can we improve our land cover dataset?*
14. What are the most common land cover classes you utilize?*
15. Would you add or expand upon any of these land cover classes?*
16. What are your frustrations with high-resolution data?
17. What do you hope to use high-resolution land cover data for?
18. Are there any tools, web apps, datasets etc., which would help you moving forward?
19. Please list any common channels or networks you use for GIS news and updates
20. Please list any listservs or data clearinghouse sites you utilize frequently
21. What are datasets you use frequently in your work other than LULC?
22. Would you be willing to be contacted about this in the future?*
23. Would you like to receive updates about our work?*
24. Do you know anyone in your network we should contact?*

Appendix B

Table 1: How respondents use land cover data in their work.

Link between land use and water quality in watersheds
We use MapShed software, which requires land use input for projecting sediment and nutrient reductions from BMP implementations
I use LULC for overland flow modeling to predict water quality in the Potomac River watershed of WV
I use impervious surface (including roads and structures) and forest cover in watershed planning including comparing water quality data to land cover
Use for analysis in various technical assistance projects with localities, for example in small area plans
Wetland functional assessment using GIS
I use land cover to evaluate conservation opportunities such as intact forest areas to conserve, agricultural land to consider for conservation opportunities, land cover vs. inundation footprints, etc. I use mainly NLCD, sometimes the PA LULC, and more recently, the UVM/CC high res dataset, as well as custom land cover datasets created through aerial photo interpretation.*
Currently upgrading to GIS datasets; comprehensive plan and capital improvements planning purposes
NLCD, MD dept of Planning LULC, UMD Tree cover/height datasets
Occasionally I need to quantify land cover percentages for a watershed or other natural area for reports
We are looking at changes and targeting areas for outreach/education
We plan to use similar data to identify where riparian buffers can be expanded to improve water quality
Socio-ecological assessment of City of Wilmington to identify areas of greatest need and opportunity for tree planting and other natural solutions projects
In partnership with the Center for Environment and Society, we use and have used land cover data to gain better understanding of land use practices in the agricultural areas of the eastern shore. We've also used it to identify best management practices for farmers and institutions interested in the sediment runoff issues
Identifying and prioritizing riparian buffer restoration sites
Most recently I have sent land cover data to college students working on projects related to our county. I am reworking a tool that I made for evaluating the value of sites of interest to birds which specifically uses land cover to identify Belted Kingfisher habitat. That is a responsibility species for New York State in their Audubon Important Bird Area Program. We use the USDA's cropland land cover map for determining agricultural use as one of the criteria for prioritizing agricultural lands for preservation.

Additionally, I use the Susquehanna River Basin Commission's land cover data and the GAP Analysis land cover data.
Modeling, deer and elk habitat suitability index, revisiting/updating land cover for our Wildlife Management Areas for use in our online hunting map, management plans, etc., watershed resources registry suitability analyses. West Virginia Terrestrial Habitat Map, LULC for WV, NLCD...interested in covering the entire state of WV with higher resolution data and more detailed classification.
We use land cover data to examine the watershed draining to our stream restoration sites. This allows us to model the hydrology and understand possible causes of impairment. We often use aerial photos combined with StreamStats to estimate watershed conditions, but true land cover data is more valuable, and I use it to check our consultants' estimates.
We use land cover data in our storm water work
Land use is intimately related to N, P, and TSS export from land to water
We use land cover data to help design and identify potential stream restoration projects. Land cover helps assess stream health, and estimate stream discharge during different rainfall events
Wetland function assessment for Clean Water Act enforcement
Tracking forest cover, buffers, urban tree canopy; targeting conservation and restoration, estimating ecosystem benefits
To explore buffer opportunities
Mapping landscape disturbance
I use land cover datasets in an academic capacity for various classroom assignments
I am currently working on a study to identify an influence between large scale watershed factors and Brook Trout abundance in the Susquehanna Watershed*

Table 2: Obstacles to using land cover data.

Our work is very project focused, so unless I get a project which focuses on or would benefit from this information I don't generally use it...Unfortunately only half our service area is in the Chesapeake Bay watershed as well, so if despite the universality of land cover data, your dataset only covers part of regions I need it to.
I have not found land cover layers that are accurate for differentiating high and low marsh, which would be my primary need. Resolution of layers is usually inadequate.
Quality of available data, time, capacity, financial resources
Staff time to develop gis-based estimates relevant to our work, capacity to do updates and change analysis for trend tracking, limitations of hydrology layer, lack of information about groundwater inputs to streams, processing capacity internally for fine-scale data at statewide scale.

Budgetary constraints, quality of available data, etc.
Time & Processing
Quality is mostly the greatest concern as far as spatial resolution and accuracy. The NLCD is almost 6 years old and is vastly out of date.
The main obstacle that I am currently experiencing with the dataset is being able to convert those data into usable, quantifiable data for my study.

Table 3: What frustrations do you have with high-resolution datasets, including but not limited to ours?

Slow loading
It seldom extends into WV! Regional products are wonderful for the surrounding states, but we often find that the data products appear to extend into WV, but the high-res drops out at the state border.
Overcoming the chicken pox effect, and continued shadow problems with classification of trees and forests
Need point cloud from canopy
Not knowing how precise they are compared to our parcel data
Dataset size and lack of storage capacity to hang onto them for extended periods of time
The sheer size of the data once it is converted makes advanced analysis difficult
The resolution of the classes and lack of forest wetland delineations is disappointing

Table 4: Are there analyses you cannot perform? If so, why?

Quality of available data, time, capacity, financial resources
Staff time
N/A
Time
Need better resolution satellite imagery
Limits to the applicability of what is requested by residents or other municipalities
Of the tools that I haven’t used, I haven’t used them due to a lack of need
I am relatively new to advanced GIS analyses and I am somewhat limited due to computational power of the computers that I have access to

Table 5: How can we improve our land cover?

It would be good if classes were consistent across state lines
Extend finer resolution mapping to WV, including wetlands and low vegetation
Consistency among states
N/A
Need to watch webinar
Include agricultural information
Utilize hyper spectral analysis to determine what type of agricultural products are being grown on particular agricultural fields
Develop vector data

Appendix C

Successful outreach and engagement in the future should be characterized by accomplishing the following:

- Informs and educates about hazards and risks
- Invites interested parties to contribute their views and ideas for mitigation
- Identifies conflicts and incorporates different perspectives and priorities early in the process
- Provides data and information that improves overall quality and accuracy of the data development plan
- Ensures transparency and builds trust
- Maximizes opportunities for implementation through greater consensus and acceptance

Thus, the following are recommendations for outreach and engagement efforts moving forward:

1) Identify and update contact list for CBPP
    a) Coordinate with user groups that are relevant, such as CBPP’s Land Use Work Group, CBPP’s Forest Work Group, to understand: 1. If data production is on both organizations agendas; 2. If not, why not?; 3. Who they talk to regularly?; 4. Same message about our data, and ensure our goal is being communicated clearly and concisely
    b) The contact list should be kept up-to-date to reflect turnover or changes in staffing
    c) Initial outreach should come from CBPP staff person because the Partnership has a “brand” and many organizations, especially in NY and northern PA, are more familiar with CBPP, so this strategy gives them more context
2) Create a master list and identify targeted groups and organizations
    a) Develop OR organize a master list by combining/labeling contacts from the involved organizations’ mailing list, mailing lists of partners associated with the land cover project, association lists (organizations in particular service areas: after-school programs etc.,), nonprofit lists, faith organizations etc.
3) Sync up postal mailing and email targeted contacts
4) Develop a distribution strategy:
    a) Post information on partners’ websites
    b) Make arrangements to post information on other websites supportive of the project
    c) Send email notices to the master email mailing list, determine a regular schedule, and leverage Constant Contact for ease
    d) Send notices by direct mail and/or fax to your master mailing list
    e) Make personal phone calls to targeted and influential groups to invite them to apply, and to ask for their help in outreach (create list of individuals known to be utilizing CBPP’s land cover data or wanted to upon interview such as, TNC, Cory Sauve, and Dr. Matthew Baker)
    f) Deliver presentations to targeted audiences at organizational meetings, conferences, trainings
    g) Create a new webinar series to discuss targeted use cases and best practices for working with raster data
    h) Place notices in publications and newsletters
    i) Bring brochures and project descriptions to conferences
    j) Curate a list of related press and post working links on the Chesapeake Conservancy website