Eastern Great Lakes 2010 Coastal Change Analysis Program Accuracy Assessment

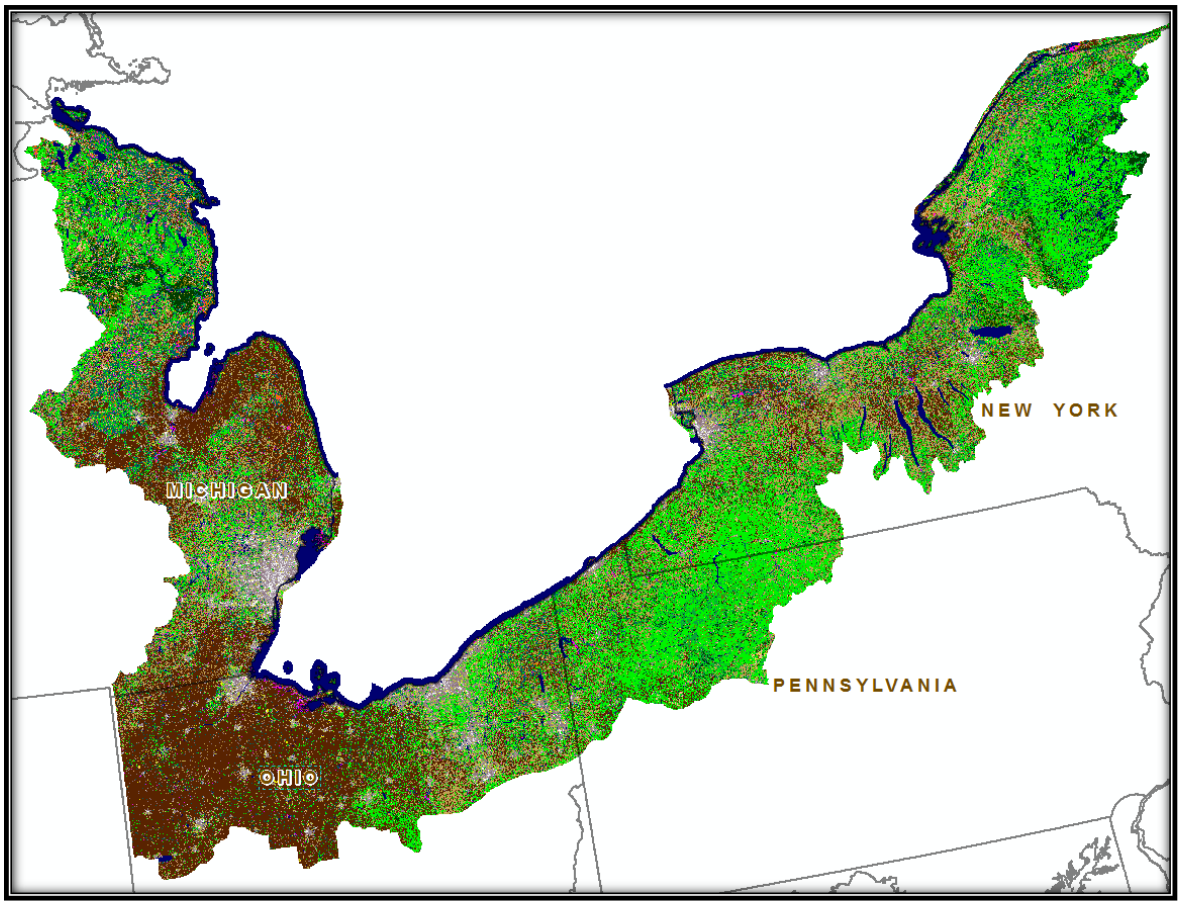
National Oceanic and Atmospheric Administration (NOAA)
Coastal Services Center

August 2014

2234 South Hobson Avenue
Charleston, South Carolina 29405-2413
(843) 740-1200
www.csc.noaa.gov

Regional Offices:
NOAA Pacific Services Center, NOAA Gulf Coast Services Center, and
Offices in the Great Lakes, Mid-Atlantic, Northeast, and West Coast

Table of Contents



Overview

This report describes the accuracy assessment that was performed on the National Oceanic and Atmospheric Administration (NOAA) 2010 Coastal Change Analysis Program (C-CAP) land cover update for the Eastern Great Lakes region. This area covers over 70,000 square miles. Before this update, the last accuracy assessment of C-CAP data for the region was performed on the 2001 baseline map product. This previous assessment was focused on the 2001 map accuracy alone and included no assessment of the change mapped. Since that time, the region has experienced a considerable amount of land cover change, and improvements have been made in detecting and mapping change. For these reasons, C-CAP determined that an accuracy assessment that included mapped change would be part of the 2010 land cover update cycle.

The 2010 Eastern Great Lakes C-CAP land cover update was conducted through the contract vehicle at the NOAA Coastal Services Center. The 2010 land cover was completed by MDA Information Systems and Sanborn. Once the external contractors completed their efforts, in-house edits were performed on all dates of land cover to address issues identified during quality assurance reviews. The C-CAP team takes extra effort to address errors in previous land cover to make a more accurate final product. Finalized land cover for the region was completed in February 2014.

Significant findings from the accuracy assessment are listed below and discussed in more detail later in this report:

- The overall accuracy for the Eastern Great Lakes 2010 C-CAP product was 84.8% (0.84 kappa).
- One class fell below 80% for both producer[1] *and* user[2] accuracy; four classes were below 80% for producer accuracy, and four were below 80% for user accuracy (Table 2).
- The accuracy for change/no-change was 86.3%, with the largest error being committed change (60.7% accuracy). It is interesting to note that of these committed change locations (falsely mapped as change) the accuracy was 83.9% for the 2010 call, indicating the 2006 call was incorrect.
- Of the 300 sample locations in mapped change areas, the accuracy was 88.3%.

Methods

The C-CAP team met and discussed accuracy assessment on multiple occasions and determined three essential requirements:

1. Ability to report overall map accuracy
2. Ability to report change/no-change mapping accuracy
3. Ability to report categorical change accuracy

A three-stratum approach (Figure 1) was chosen, including (1) current change, (2) near current and recent change, and (3) the remaining area. Stratum 1 (red) was the 2006-2010 mapped change areas. The team wanted to sample enough locations within currently mapped change to be able to assess the quality of the newly mapped areas, as well as comment on the change/no-change mapping accuracy. The team

[1] Related to errors of omission when an area is excluded from the category to which it belongs.
[2] Related to errors of commission when an area is included incorrectly in a category.

attempted to split the non-change area evenly into the other two strata. Stratum 2 (purple plus gray) was determined by combining all changes from 1996-2010 and buffering until the area target was approximated, in this case a 19-pixel buffer. This second stratum *did not* resample Stratum 1. From past experience team members have noticed that change is often spatially auto-correlated, which means that new change tends to occur near previous change. This can easily be seen in urban expansion or in the clustering of timber activity. The team felt that sample units in this stratum may be useful in potentially identifying missed change, as well as be used for wall-to-wall accuracy. The remaining area was Stratum 3 (blue). These points may pick up missed change but would be most useful in assessing wall-to-wall accuracy. Each stratum contained 300 accuracy assessment sample units.

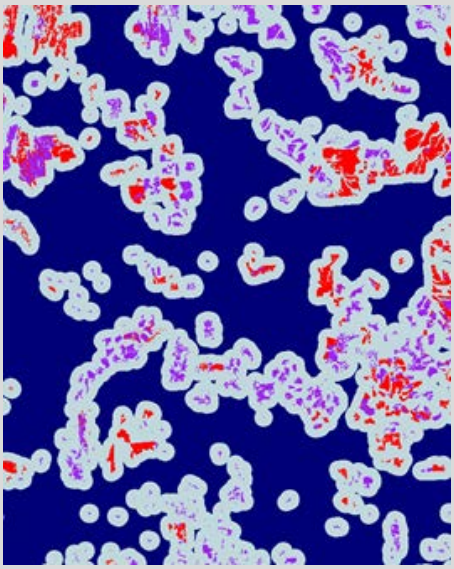


Figure 1: Accuracy Assessment Strata

Sample units were identified using the ERDAS Imagine Accuracy Assessment tool. A total of 300 sample units per stratum (total 900) were placed with the following criteria: stratified random placement; a minimum of 10 per class (not always met); and six out of nine land cover pixels around the location required to be homogenous, or else the location was discarded. The sample locations were then buffered by 45 meters to assist in interpretation of the appropriate land cover and change call.

Sample Unit Distribution and Interpretation

As seen in Table 1, Deciduous Forest and Scrub/Shrub received the most accuracy assessment sample units (97) and Palustrine Aquatic Bed received the fewest (4). The last two columns in Table 1 can be compared to assess if a class was sampled proportionally to the area it comprised. For example, Mixed Forest received 3.9% of the accuracy assessment (AA) sample units and comprised 3.1% of the region. The largest discrepancy is with Cultivated Crops receiving 10.2% of the AA sample units and comprising 25.3% of the region. Discrepancies may be due to rarer classes and classes commonly associated with change/transition, such as Bare Land.

The AA sample units were randomly split into three groups of 600 points. Each reviewer (three total) was responsible for labeling the AA sample unit according to its primary land cover using the available Landsat imagery (2010), a “fuzzy call” if necessary, and whether the sample unit changed from 2006. Fuzzy calls were used if the interpreter could not positively identify a single dominant land cover (e.g., natural speckling of land cover classes) or when land classes were very similar (e.g., Shrub vs. Forest are distinguished by a height criteria). Reviewers had access to all 2006 and 2010 Landsat data, Google Earth, National Wetlands Inventory (NWI), Soil Survey Geographic (SSURGO) database, and other high-resolution imagery (e.g., Bing Maps) as available. All points were compiled into a single file for comparison of land cover and change calls. Any locations where the review calls differed were separated for further discussion by the reviewers and project lead, if needed.

The land cover and change determinations, or "calls," for a 3 x 3 pixel window at each AA location were extracted from the data to compare against the reviewer calls. To be labeled "correct," six out of the nine map pixels had to match the primary or fuzzy review call (for land cover or change/no-change).

Table 1. Breakdown of accuracy assessment sample units per strata and per land cover class.

	Accuracy Assessment Sample Units				Percent of	
Land Cover	Stratum 1	Stratum 2	Stratum 3	Total	Sample Units	Region
Developed, High Intensity	16	12	11	39	4.3%	0.6%
Developed, Medium Intensity	21	15	12	48	5.3%	1.6%
Developed, Low Intensity	27	16	13	56	6.2%	4.5%
Developed, Open Space	31	13	11	55	6.1%	2.2%
Cultivated Crops	12	34	46	92	10.2%	25.3%
Pasture/Hay	8	23	37	68	7.6%	11.3%
Grassland/Herbaceous	28	17	12	57	6.3%	1.9%
Deciduous Forest	20	38	39	97	10.8%	24.8%
Evergreen Forest	1	15	15	31	3.4%	4.4%
Mixed Forest	1	20	14	35	3.9%	3.1%
Scrub/Shrub	61	20	16	97	10.8%	3.1%
Palustrine Forested Wetland	3	20	18	41	4.6%	7.6%
Palustrine Scrub/Shrub Wetland	18	15	13	46	5.1%	1.5%
Palustrine Emergent Wetland	17	15	12	44	4.9%	1.2%
Unconsolidated Shore	2	3	0	5	0.6%	0.0%
Bare Land	16	11	7	34	3.8%	0.3%
Open Water	14	13	24	51	5.7%	6.6%
Palustrine Aquatic Bed	4	0	0	4	0.4%	0.0%
Total	300	300	300	900		
Area (square miles)	752	36,880	32,556	70,188		
Percent of Region	1.1%	52.5%	46.4%			

Results and Discussion

2010 Land Cover

Table 2 represents the error matrix for the 2010 land cover map. Overall accuracy for the 2010 land cover product was 84.8% (0.84 kappa). The majority of classes met the C-CAP target specification of 80% per class accuracy. Of the eight instances where accuracy was below the targeted 80%, five of these did exceed 70%. One class, Grassland, fell below the 80% threshold for both producer and user accuracy. Grassland was confused with other low vegetation classes, including Open Space Developed, Cultivated Crops, and Pasture/Hay. Bare Land was confused with Pasture/Hay, Scrub/Shrub, and developed classes. The class with the lowest single accuracy was Grassland (64.7% user accuracy). The accuracies associated with

Unconsolidated Shore and Palustrine Aquatic Bed should be considered with care, as both classes had very few accuracy assessment locations (Table 2).

Although most classes did not have more than 50 sample units (the coarse "rule-of-thumb" for accuracy assessment), eight classes did exceed 50, and 12 classes were over 40. As mentioned, Unconsolidated Shore and Palustrine Aquatic Bed had few assessment locations, which can be explained by the limited area these classes represent in the landscape.

Few clear patterns were found during the examination of the off-diagonal values in the error matrix. Closer examination of the errors did reveal some trends in the classification.

1. **Low vegetation confusion** – Grassland was the only class that did not meet the 80% individual class accuracy for both user and producer accuracy. There was confusion among Grassland, Pasture/Hay, Cultivated Crops, and Open Space Developed. All of these classes often have very similar spectral characteristics within Landsat imagery, which leads to this confusion.
2. **Deciduous** – Deciduous forest was the dominant forest type in the region and is commonly associated with silvicultural activities. As forests are cut, a cycle of Grassland to Scrub/Shrub to Forest is often seen. C-CAP distinguishes between Scrub/Shrub through a height criterion, which cannot be reliably determined in Landsat imagery. This difficulty in determining vegetation height, combined with the spectral similarities of leaf-off deciduous forests and Grassland, leads to confusion among these classes.
3. **Mixed Forest** – Mixed forest occurs when Deciduous and Evergreen forest both occur and neither exceeds 75% of the total tree cover. Accurately estimating this threshold from Landsat is difficult. Additionally, large extents of forestland in Pennsylvania and New York have a deciduous canopy with an evergreen understory. When using leaf-on imagery, the forest appears deciduous, when using leaf-off imagery the forest appears evergreen. The spectral change over time in Landsat is similar to a mixed forest. More thorough examination of higher spatial resolution data or extensive field data is needed to more accurately map mixed forest.
4. **Scrub/Shrub** – Scrub/Shrub was confused with many classes, most commonly with upland forest classes. This error was also seen in other regions. As mentioned above, Scrub/Shrub is generally a transitional class between Grassland and Forest classes distinguished by height. Since height cannot be directly measured in the Landsat data used, other criteria must be used (tone, texture, shadow, etc.), resulting in the confused classes.

Table 2. Full error matrix for the 2010 Eastern Great Lakes C-CAP mapping region. Map classes are along the left edge, and reference calls are along the top of the matrix. Correct locations are highlighted in green along the diagonal of the matrix. Individual class accuracies that fall below the target 80% are highlighted in orange.

Map \ Reference	Developed, High Intensity	Developed, Medium Intensity	Developed, Low Intensity	Developed, Open Space	Cultivated Crops	Pasture/Hay	Grassland/Herbaceous	Deciduous Forest	Evergreen Forest	Mixed Forest	Scrub/Shrub	Palustrine Forested Wetland	Palustrine Scrub/Shrub Wetland	Palustrine Emergent Wetland	Unconsolidated Shore	Bare Land	Open Water	Palustrine Aquatic Bed	Total	Users
Developed, High Intensity	39																		39	100.0%
Developed, Medium Intensity	4	43									1								48	89.6%
Developed, Low Intensity		1	49	4			1							1					56	87.5%
Developed, Open Space				46	1		3	2			1		1			1			55	83.6%
Cultivated Crops			1	2	80	2	3				3		1						92	87.0%
Pasture/Hay			2	4	7	44	5	1			4			1					68	64.7%
Grassland/Herbaceous				2	3	2	40	4	1		3			1			1		57	70.2%
Deciduous Forest							1	93			1		1	1					97	95.9%
Evergreen Forest									28	3									31	90.3%
Mixed Forest								3		32									35	91.4%
Scrub/Shrub			1	3			2	15	2	4	66	3		1					97	68.0%
Palustrine Forested Wetland												40	1						41	97.6%
Palustrine Scrub/Shrub Wetland											1	5	39	1					46	84.8%
Palustrine Emergent Wetland												1	5	35			1	2	44	79.5%
Unconsolidated Shore															5				5	100.0%
Bare Land				1	1		1		1							30			34	88.2%
Open Water			1														50		51	98.0%
Palustrine Aquatic Bed																		4	4	100.0%
Total	44	54	62	92	48	56	118	32	39	80	49	48	41	5	31	52	6		900	
Producers	90.7%	97.7%	90.7%	74.2%	87.0%	91.7%	71.4%	78.8%	87.5%	82.1%	82.5%	81.6%	81.3%	85.4%	100.0%	96.8%	96.2%	66.7%		84.8%

Fuzzy calls were allowed in conditions where the field class was either difficult to positively identify (e.g., Cultivated vs. Pasture, Shrub vs. Forest, different levels of development) or where there was natural variability in the landscape (e.g., near edge features). Using fuzzy calls increases the chance for a correct label but may potentially artificially inflate the reported map accuracy if they are overused. Table 3 shows that although 29% of the sample units received a fuzzy call, these calls were rarely responsible (13%) for a location being deemed mapped as correct.

Table 3. Fuzzy reference calls for the 2010 Eastern Great Lakes C-CAP region.

Fuzzy Reference Calls		
Of the 900 sample locations, 259 (28.8%) had a fuzzy call	For the 763 correctly mapped locations, 98 (12.8%) were correct based on the fuzzy land cover call (665 were correct based on primary call)	Land cover classes with the most fuzzy calls include different levels of Development, Scrub and Grass, and Palustrine Forest and Palustrine Scrub/Shrub

2006-2010 Change

Overall change/no-change accuracy was 86% (Table 4). Committed change was the largest error, with a user accuracy of 61% (118 sample locations mapped as change, but deemed no change by the reviewers). These 118 locations were assessed in their own error matrix and resulted in 83.9% overall accuracy. This seems to indicate that the method used to identify potential change pixels (creating the change mask) may be overestimating change, but the methods used to assign a land cover class are reasonably accurate. These locations of committed change may be used in future editing efforts, since they are indicative of potential errors with the 2006 map.

Assessing mapped change is a fairly straightforward task, but assessing missed change is problematic. Of the 600 total sample units in mapped no-change areas, only five were deemed missed change. Four of these points were within Stratum 2 (specifically designed to try to identify potential missed change). After conducting the change analysis, the team feels that overall change has been slightly overcalled, with limited missed true change.

Table 4. Change/no-change matrix for the 2010 Eastern Great Lakes C-CAP region. Correct locations are highlighted in green along the diagonal of the matrix. Change calls were coded *0* for no change, and *1* for change.

		Reference Change			
		0	1	Total	Users
Map Change	0	595	5	600	99%
	1	118	182	300	61%
	Total	713	187	900	
	Producers	83%	97%		86%

A final analysis was performed using only sample locations interpreted as change (187 locations). Table 5 shows that the overall accuracy of these locations was 89.8%, slightly higher than the total map accuracy (Table 2). Errors within this matrix were similar to the overall matrix discussed previously.

Comparison between 2001 and 2010 Accuracy Assessments

This 2010 accuracy assessment was covered by portions of two accuracy assessments for the 2001 products (Great Lakes and Michigan). The 2001 Great Lakes report covered portions of New York, Pennsylvania, Ohio, Indiana, Illinois, Wisconsin, and Minnesota and had an overall reported accuracy of 91.4%. The 2001 Michigan assessment reported an 87.7% overall accuracy. As stated above, this 2010 accuracy assessment reported an overall 84.8% accuracy. It is important to highlight potential reasons for these differences.

Report	Coverage	Land Cover Classes	AA site criteria (3x3 window)	Collection method	Sample units
2001	NY, PA, OH, IN, IL, WI, MN, MI (152.667 mi^2)	15	all 9 land cover pixels had to be homogenous	in the field, driving, aerial	5,941
2010	OH, portions of MI, PA, NY (70,188 mi^2)	18	6 out of 9 agree	photo-interpretation, Google Earth, NAIP, NWI, Landsat ancillary data	900

Accuracy assessment locations for the 2001 product were photointerpreted from black and white digital orthophotography and footprints of high-resolution satellite imagery. All sample locations were drawn from a 3 x 3 pixel window of homogenous land cover. Sampling from within imagery footprints can be seen in the clustering of sample locations in Figure 2. The restriction of sample locations to a homogenous area reduced the chance of speckled or more-difficult-to-classify areas being chosen for accuracy assessment.

Although the reported accuracy for the 2001 land cover product is higher than the 2010 product, the 2010 accuracy number appears to be more representative of the true quality of the map for several reasons.

- Autocorrelation in the data: Field-collected data were heavily used in the 2001 project. Data were generally collected along roadways while transiting the region. Accuracy assessment locations (Figure 2) were then sampled from this data set before image classification. Often, multiple accuracy assessment locations would fall within a single feature. Sampling along roadways also may introduce bias because roads are typically located to avoid certain features (e.g., wetlands, mountains, poor soils for construction), and thus land cover types in these locations may not be sampled. Limited aerial surveys were performed in an attempt to sample away from roads. Accuracy assessment locations for the 2010 map were randomly placed in one of three strata. Although there may be some bias with the 2010 data (30% of the sample units were placed in Stratum 1 which comprised 1.1% of the area), the overall distribution appears much more representative of the region (Figure 3).
- Number of classes: The 2001 land cover map had three fewer land cover classes to map (no Developed Medium Intensity, Developed Open Space, or Pasture/Hay).
- Filtering of land cover: The accuracy assessment locations for the 2001 project were isolated to 3 x 3 homogenous areas of land cover. The 2010 accuracy assessment allowed some of the natural variability in the land cover to be included in the map (six out of nine pixel homogeneity in a 3 x 3 window).
- In-house editing of final land cover: The C-CAP team took considerable effort in addressing identified issues with the previous dates of land cover. Many of these fixes were consistent through time, thus improving all dates of land cover.

Table 5. Error matrix for the 2010 Eastern Great Lakes C-CAP mapping region based on interpreted change locations. Map classes are along the left edge and reference calls are along the top of the matrix.

Map \ Reference	Developed, High Intensity	Developed, Medium Intensity	Developed, Low Intensity	Developed, Open Space	Cultivated Crops	Pasture/Hay	Grassland/Herbaceous	Deciduous Forest	Evergreen Forest	Scrub/Shrub	Palustrine Forested Wetland	Palustrine Scrub/Shrub Wetland	Palustrine Emergent Wetland	Unconsolidated Shore	Bare Land	Open Water	Palustrine Aquatic Bed	Total
Developed, High Intensity	13																	13
Developed, Medium Intensity		14								1								15
Developed, Low Intensity			11	1														12
Developed, Open Space				4				1		1								6
Cultivated Crops					3	1	1			1								6
Pasture/Hay						2												2
Grassland/Herbaceous							21			1								22
Deciduous Forest							1	7		1								9
Scrub/Shrub							2		2	50	1							55
Palustrine Scrub/Shrub Wetland											1	10	1					12
Palustrine Emergent Wetland													8				1	9
Unconsolidated Shore														2				2
Bare Land					1										12			13
Open Water																9		9
Palustrine Aquatic Bed																	2	2
Total	13	14	11	5	4	3	25	8	2	55	2	10	9	2	12	9	3	187

Correct	168
Percent Correct	89.8%

Figure 2. Accuracy assessment locations for the 2001 C-CAP Great Lakes area draped on 2001 land cover. Two accuracy assessments were performed for this region and are displayed as separate colors (red for Michigan, black for the remaining area of the Great Lakes C-CAP region). Sampling along road networks can be seen, along with clustering of sample locations (e.g., Grand Rapids, Minnesota, in the northwestern portion of the map).

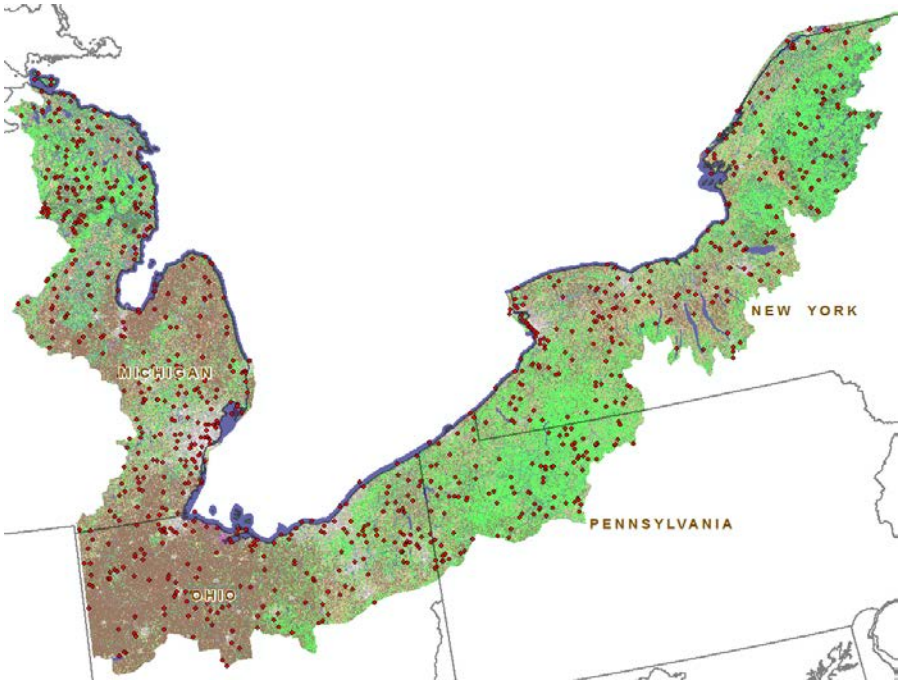


Figure 3. Accuracy assessment locations for the 2010 C-CAP Eastern Great Lakes area draped over the 2010 land cover.

Conclusions

C-CAP uses consistent methods and approaches for mapping land cover and land cover change for the coastal regions of the U.S. with a stated accuracy target of 85% overall and 80% per class. The 2010 Eastern Great Lakes region was assessed for accuracy through in-house efforts. Sampling strata were established to estimate overall accuracy, as well as change mapping accuracy. The overall accuracy of the region was 84.8%, with the majority of individual classes exceeding 80% accuracy. Change/no-change accuracy for the product was 86.3%, with committed change being the largest error. The analysis found that 83.9% of the false change locations received the correct 2010 call, indicating the classification approaches appear to be working well.

Only one class, Grassland, had accuracy below 80% for both user and producer accuracy. The confusion of this class and other low vegetation classes has been seen in previously mapped regions and discussed earlier. Addressing this error will most likely involve creating better developed surface masks, to help eliminate Open Space Developed, and collaborating with the National Agricultural Statistics Service (NASS) to improve Cultivate Crops and Pasture/Hay mapping.

By its definition, Mixed Forest is a confused class; it is made from portions of two other forest types. The classification of this class can become more difficult based on the phenological condition of the imagery used (leaf on in the growing season, leaf off during the winter). As discussed earlier, there are areas in this region with a deciduous canopy and evergreen understory. This combination of land cover, which should be classified as Deciduous Forest, tends to have the same spectral appearance over time as Mixed Forest when viewed in Landsat imagery. More research will be needed to determine the best approach to correcting this issue.

Confusion among Grassland, Scrub/Shrub, and Forest remains a difficult issue to address. These classes are most commonly found in areas with timber activity. Separating Scrub/Shrub from Forest is generally based on a height criterion, which cannot be directly measured from Landsat. As lidar data become more available in the future, these data may be used to estimate vegetation height and help separate Scrub/Shrub from Forest.