Western Great Lakes 2010 Coastal Change Analysis Program Accuracy Assessment

National Oceanic and Atmospheric Administration (NOAA)
Coastal Services Center

June 2014

NOAA Coastal Services Center

LINKING PEOPLE, INFORMATION, AND TECHNOLOGY

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 1

Methods 1

Sample Unit Distribution and Interpretation 2

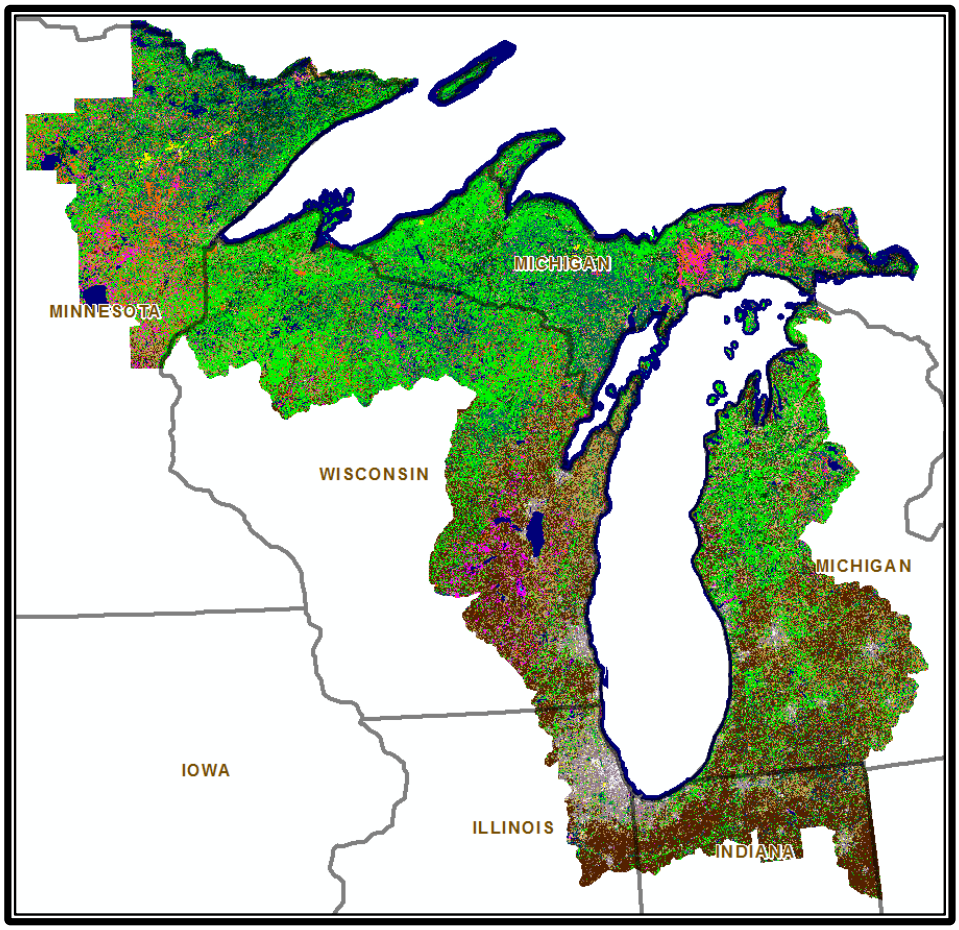
Results and Discussion 3

2010 Land Cover 3

2006-2010 Change 6

Differences between 2001 and 2010 Accuracy Assessments 7

Conclusions 11



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 Western Great Lakes. This area covers over 107,000 square miles and includes the coastal portions of Minnesota, Wisconsin, Illinois, Indiana, and Western Michigan. 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, there has been both a considerable amount of land cover change in the region and improvements in detecting and mapping change. For these reasons, it was determined that an accuracy assessment, which included mapped change, would be part of the 2010 land cover update cycle.

The 2010 Western Great Lakes C-CAP land cover update was made up of two task orders (Lake Michigan/Wisconsin and Lake Superior) through the contract vehicle at the NOAA Coastal Services Center. The 2010 land cover was completed by Photo Science. Once the external contractor completed its 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 late December 2012.

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

- The overall accuracy for the Western Great Lakes 2010 C-CAP product was 85.6% (0.84 kappa).
- No class fell below 80% for both producer[1] *and* user[2] accuracy; two classes were below 80% for producer accuracy, and four were below 80% for user accuracy (Table 2).
- The accuracy for change/no change was 88.7%, with the largest error being committed change (68.0% accuracy). It is interesting to note that of these committed change locations (falsely mapped as change) the accuracy was 80.2% for the 2010 call, indicating the 2006 call was incorrect.
- Of the 300 sample locations in mapped change areas, the accuracy was 85.7%.

Methods

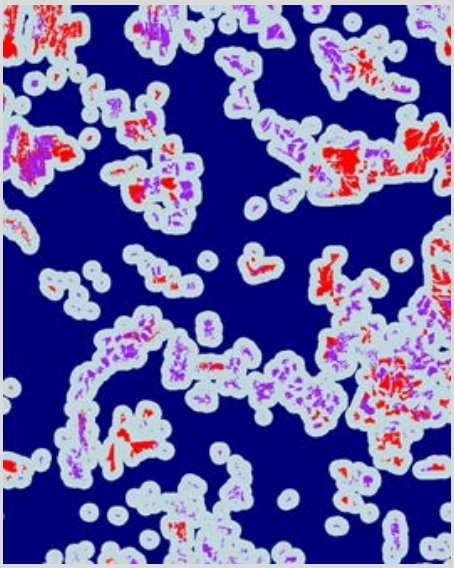
The C-CAP team met and discussed accuracy assessment on multiple occasions and determined three essential requirements:

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

[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.

newly mapped areas, as well as comment on the change/no-change mapping accuracy. The team 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 14 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.



Accuracy Assessment Strata
Figure 1: Accuracy Assessment Strata

Sample units were identified using the ERDAS Image 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 had 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 received the most accuracy assessment sample units (123) and Unconsolidated Shore received the least (9). 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 4.7% of the accuracy assessment (AA) sample units and comprised 5.6% of the region. The largest discrepancy is with Deciduous Forest receiving 13.7% of the AA sample units and comprising 21.3% 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	9	11	10	30	3.3%	0.5%
Developed, Medium Intensity	5	12	11	28	3.1%	1.2%
Developed, Low Intensity	18	16	14	48	5.3%	3.2%
Developed, Open Space	18	12	11	41	4.6%	1.1%
Cultivated Crops	13	29	37	79	8.8%	16.0%
Pasture/Hay	5	19	19	42	4.8%	6.3%
Grassland/Herbaceous	49	12	12	72	8.0%	2.5%
Deciduous Forest	41	40	40	123	13.7%	21.6%
Evergreen Forest	14	19	19	46	5.1%	5.4%
Mixed Forest	7	18	18	42	4.7%	5.6%
Scrub/Shrub	55	16	16	81	9.0%	3.8%
Palustrine Forested Wetland	12	27	27	66	7.3%	12.6%
Palustrine Scrub/Shrub Wetland	17	21	18	56	6.2%	5.9%
Palustrine Emergent Wetland	15	16	13	43	4.8%	3.2%
Unconsolidated Shore		6	3	9	1.0%	0.0%
Bare Land	13	11	10	34	3.8%	0.4%
Open Water	9	16	35	60	6.7%	10.6%
Total	300	300	300	900		
Area (square miles)	2,220	52,522	52,877	107,619		
Percent of Region	2.1%	48.8%	49.1%			

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 85.6% (0.84 kappa). Nearly all classes met the C-CAP target specification of 80% per class accuracy. Two classes, Evergreen Forest and Palustrine Forest, fell below the 80% threshold for producer accuracy. Four classes (Developed Low Intensity, Pasture/Hay, Palustrine Scrub/Shrub, and Unconsolidated Shore) fell below 80% for user accuracy. Pasture/Hay had the lowest single class accuracy (64.33% user accuracy) and was most confused with Cultivated.

Although most classes did not have more than 50 sample units (the coarse “rule-of-thumb” for accuracy assessment), seven classes did exceed 50, and 12 classes were over 40. The fewest sample units were associated with Unconsolidated Shore.

Table 2. Full error matrix for the 2010 Western 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	Total	Users
Developed, High Intensity	26	2					1									1		30	86.7%
Developed, Medium Intensity	1	25	1	1														28	89.3%
Developed, Low Intensity		1	37	2	4	1	2										1	48	77.1%
Developed, Open Space				37		1	1	1								1		41	90.2%
Cultivated Crops		1		1	68	3	1	1	1					2		1		79	86.1%
Pasture/Hay			1	3	7	27	2		1	1								42	64.3%
Grassland/Herbaceous			1	1	2		63	2	1		1	1						72	87.5%
Deciduous Forest							2	107	4		5	3		2				123	87.0%
Evergreen Forest									42		2	2						46	91.3%
Mixed Forest								1	2	37	1	1						42	88.1%
Scrub/Shrub							1	7	1	5	65		2					81	80.2%
Palustrine Forested Wetland									1			64	1					66	97.0%
Palustrine Scrub/Shrub Wetland				1				1			1	10	43					56	76.8%
Palustrine Emergent Wetland					1	1	1					2	2	36				43	83.7%
Unconsolidated Shore														2	7			9	77.8%
Bare Land					1		1		1					1		29	1	34	85.3%
Open Water								1		1				1			57	60	95.0%
Total	27	29	40	46	83	33	75	121	54	44	75	83	48	44	7	32	59	900	
Producers	96.3%	86.2%	92.5%	80.4%	81.9%	81.8%	84.0%	88.4%	77.8%	84.1%	86.7%	77.1%	89.6%	81.8%	100.0%	90.6%	100.0%		85.6%

There were two major sources of classification confusion as seen in the error matrix:

1. **Cultivated and Pasture/Hay** – These two classes are very hard to separate spectrally using Landsat data. Higher resolution data, more dates of spectral data (to capture phenology), and more field data are often needed to more accurately separate these classes.
2. **Scrub/Shrub being classified as Grassland or Forest classes, and vice versa** – This occurred in both upland and wetland conditions. All these classes are often directly related to each other, as one class transitions to the next through forest cutting and regeneration. The Scrub/Shrub class is generally a transitional class between Grassland and Forest classes and is distinguished in C-CAP by a height criterion. 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.

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 53% of the sample units received a fuzzy call, these calls were rarely responsible (17%) for a location being deemed mapped as correct.

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

Fuzzy Reference Calls		
Of the 900 sample locations, 476 (52.8%) had a fuzzy call	For the 770 correctly mapped locations, 135 (17.5%) were correct based on the fuzzy land cover call (635 were correct based on primary call)	Land cover classes with most fuzzy calls include Grass–Scrub/Shrub, Palustrine Scrub/Shrub–Palustrine Emergent, and Developed Open Space–Developed Low Intensity

2006-2010 Change

Overall change/no change accuracy was 89% (Table 4). Committed change was the largest error with a user accuracy of 68% (96 sample locations mapped as change, but deemed no change by the reviewers). These 96 locations were assessed in their own error matrix and resulted in an 80.2% 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 fairly 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 six were deemed missed change. Four of these points were within Stratum 2 (specifically designed to try to identify

potential missed change). Based on the change analysis, the team feels that mapped change may be slightly overcalled (although this committed change is generally the correct 2010 call) and that there is very little omitted change in the map.

Table 4. Change-no change matrix for the 2010 Western 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	594	6	600	99%
	1	96	204	300	68%
	Total	690	210	900	
	Producers	86%	97%		89%

A final analysis was performed using only sample locations interpreted as change (210 locations). Table 5 shows that the overall accuracy of these locations was 87.1%, similar to the total map accuracy (Table 2). The Scrub/Shrub class appeared to have the most confusion, being mapped as other natural vegetation classes (Grassland and Forest, both upland and wetland). This confusion is similar to what was seen in Table 2.

Differences 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 85.6% 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	MN, WI, IL, IN, portion of MI (107,619 mi^2)	18	6 out of 9 agree	photo-interpretation, Google Earth, NAIP, NWI, Landsat, ancillary data	900

Table 5. Error matrix for the 2010 Western 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	Mixed Forest	Scrub/Shrub	Palustrine Forested Wetland	Palustrine Scrub/Shrub Wetland	Palustrine Emergent Wetland	Bare Land	Open Water	Total
Developed, High Intensity	6	2					1										9
Developed, Medium Intensity	1	2	1														4
Developed, Low Intensity			6				1									1	8
Developed, Open Space				5		1									1		7
Cultivated Crops					5										1		6
Pasture/Hay									1								1
Grassland/Herbaceous							43				1						44
Deciduous Forest								13	2		3			1			19
Evergreen Forest									9		1						10
Mixed Forest										4	1						5
Scrub/Shrub							1	2		1	47		1				52
Palustrine Forested Wetland												3					3
Palustrine Scrub/Shrub Wetland													14				14
Palustrine Emergent Wetland														12			12
Bare Land					1										8		9
Open Water								1								6	7
Total	7	4	7	5	6	1	46	16	12	5	53	3	15	13	10	7	210

Correct	183
Percent Correct	87.1%

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 2.1% 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.

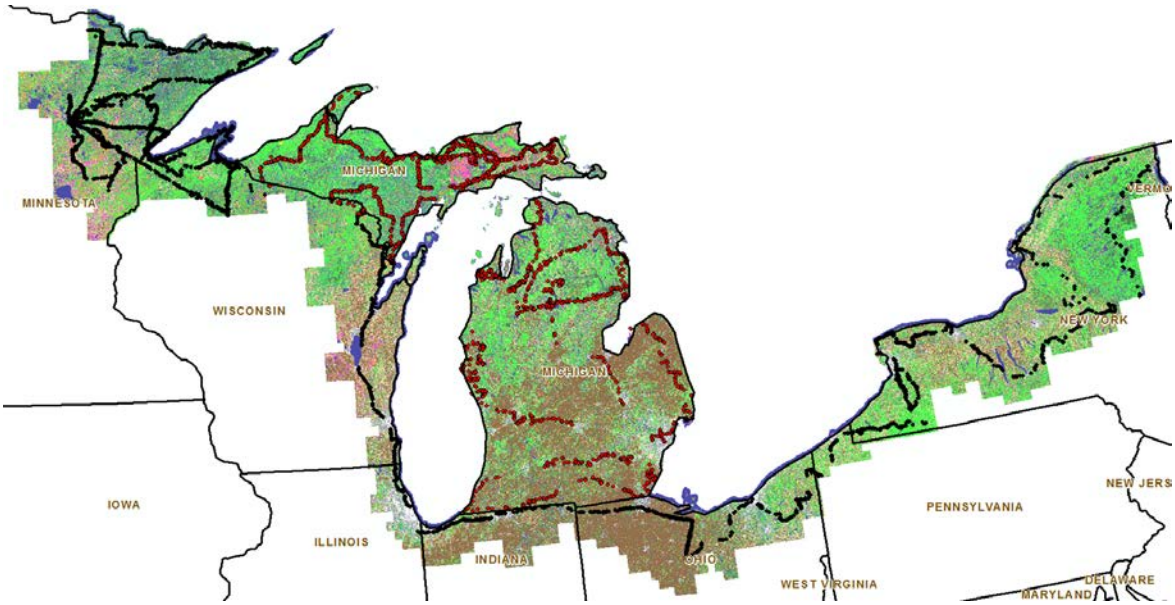


Figure 2. Accuracy assessment locations for the 2001 C-CAP Great Lakes area draped on 2001 land cover. Two accuracy assessments were performed for this region and are displayed as separate colors (red for Michigan, black for the remaining area of the Great Lakes C-CAP region). Sampling along road networks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

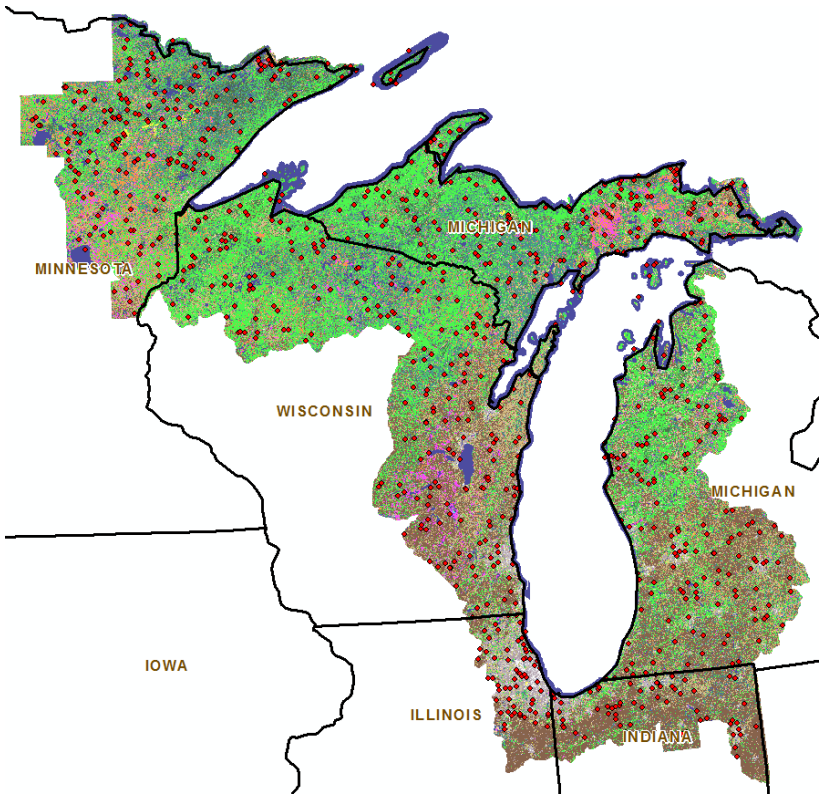


Figure 3. Accuracy assessment locations for the 2010 C-CAP Great Lakes area draped on the 2010 land cover.

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 Western Great Lakes region was the first C-CAP area that 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 85.6%, with nearly all individual classes exceeding 80% accuracy. There were very few trends to be found in the error matrix outside of confusion between Cultivated/Pasture and Grass/Shrub/Forest classes. Change/no-change accuracy for the product was 89%, with committed change being the largest error. It was found that 80% of the false change locations received the correct 2010 call, indicating the classification approaches appear to be working well.

Although the accuracy was high, potential improvements could be performed on the map in the future, or as part of the next update cycle. The Cultivated vs. Pasture/Hay classes may be improved through future work with the National Agricultural Statistics Service (NASS). As NASS develops and improves its crop-mapping approaches, its data may be incorporated into C-CAP. Finally, the U.S. Geological Survey is working on improving its impervious surface mapping for both the current time frame and back dates. These data may be used to help refine the Developed categories.