Accuracy Assessment of NOAA Coastal Change Analysis Program 2006-2010 Land Cover and Land Cover Change Data

John W. McCombs, Nathaniel D. Herold, Shan G. Burkhalter, and Christopher J. Robinson

Abstract

A new approach to locating accuracy assessment sample units was used to quantify 2010 land cover accuracy, in addition to being able to make statements about 2006-2010 land cover change mapping accuracy for National Oceanic and Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP) data. Three customized tiers of sampling strata were created, as discussed, to meet these goals. Stratified random sampling was employed in each stratum with a six out of nine pixel-homogeneity criteria (different from the final minimum mapping unit) required for each sampling unit. Accuracy was assessed for nine regions in the coastal United States with overall accuracy ranging from 82.3 percent to 85.6 percent. Binary change was mapped with 88.7 percent accuracy, with the largest error being errors of commission (71.2 percent user accuracy). This sampling design also allowed for the identification of 137 locations where true change was not mapped, allowing for statements to be made about missed change.

Introduction

Land cover and land cover change are of critical importance, with implications for water quality (Kang *et al.,* 2014; Lu and Weng, 2006; Margriter *et al.,* 2014), wildlife habitat (Lowe and Peterson, 2014; Millette *et al.,* 2014; Porter *et al.,* 2015), forest fragmentation (Civco *et al.,* 2002; Nagendra *et al.,* 2004), ecosystem health (Greene *et al.,* 2014; Lowe and Peterson, 2014; Nestlerode *et al.,* 2014), human health (Cleckner and Allen, 2014; Liang and Gong, 2015; Raghavan *et al.,* 2014), and climate change (Galbraith *et al.,* 2002; Hansen and Loveland, 2012; Morris *et al.,* 2002).

In 2010, over 123 million people, or 39 percent of the nation's population, lived in coastal shoreline counties, representing less than 10 percent of the U.S. land area (excluding Alaska) (U.S. Census Bureau, 2011). Population density within this area is expected to increase to 14.26 persons per square kilometer (37 persons per square mile) by 2020, while the expected increase for the entire U.S. is 4.25 persons per square kilometer (11 persons per square mile) (Woods and Poole Economics, 2011). Economic data for the U.S. indicate that ocean- and Great Lakes-dependent businesses employed 2.8 million people, paid $107.5 billion in wages, and produced $282.2 billion in goods and services in 2011. From 2010 to 2011, the ocean and Great Lakes economy gained 67,000 jobs, an increase of 2.4 percent - twice the employment growth rate of the U.S. economy as a whole. Real gross domestic product grew by 2.7 percent, faster than the U.S. economy as a whole (1.6 percent) (NOAA ENOW). Because of the importance of the U.S. coastal zones, and the rate at which change is occurring within it, the need for accurate and timely land cover and land cover change data is vital.

Through its Coastal Change Analysis Program (C-CAP), the National Oceanic and Atmospheric Administration (NOAA) produces land cover for the coastal regions of the United States. C-CAP inventories coastal intertidal areas, wetlands, and adjacent uplands with a goal of monitoring these habitats by updating the land cover every five years (Dobson *et al.,* 1985; NOAA C-CAP, 2015). Resulting data are then incorporated into, and serve as the coastal expression of, the United States Geological Survey (USGS) National Land Cover Database (NLCD).

The nationwide C-CAP baseline was developed from imagery acquired by the Landsat suite of satellites, *circa* 2001, using a standardized classification approach (unsupervised/supervised classification, spatial modeling, hand edits, etc.). Since that time, additional dates have been created for 1996, 2006, and 2010, in addition to limited geographies having data for 1985 and 1992. The C-CAP approach to creating a new date of land cover consists of identifying potential areas of change between multiple dates of Landsat data (e.g., 2006 to 2010) through a variety of spectral change analyses, classifying those areas of potential change in the new date, then overlaying classified areas of change over remaining non-change areas to create a wall-to-wall map for the new date. Over time, as additional dates of land cover were created, the change detection and mapping methods have improved, previous errors have been addressed, and significant steps have been taken to improve the overall quality of the map. The completion of the 2010 data incorporated the largest improvements to date, including improved impervious surface/developed land cover from the USGS and improved wetland/upland distinction through the development and application of a NOAA Office for Coastal Management wetland potential layer.

A previous accuracy assessment of C-CAP land cover was performed on the 2001 data set. This first assessment focused on the single-date map accuracy and included no assessment of the mapped change (1996 to 2001). Since that time, new land cover classes have been added, the nation 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 of the 2010 wall-to-wall land cover and mapped 2006-2010 change would be part of this update cycle.

John W. McCombs, Shan G. Burkhalter, and Christopher J. Robinson are with The Baldwin Group at NOAA Office for Coastal Management, 2234 South Hobson Avenue, Charleston, SC 29405 (john.mccombs@noaa.gov).

Nathaniel D. Herold is with the NOAA Office for Coastal Management, 2234 South Hobson Avenue, Charleston, SC 29405.

Photogrammetric Engineering & Remote Sensing
Vol. 82, No. 9, September 2016, pp. 711–718.
0099-1112/16/711–718

doi: 10.14358/PERS.82.9.711

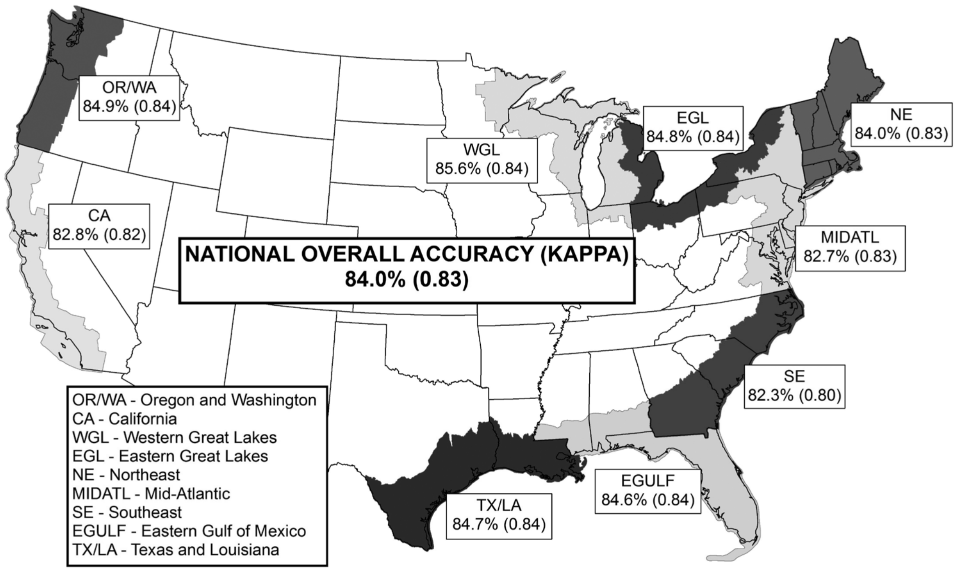

Figure 1. The C-CAP mapping area was divided into nine regional assessments. The overall accuracy and kappa are indicated for each region and for the national level.

Methods

The accuracy assessment of remotely sensed products (e.g., land cover and change maps) has three major components that should be decided upon before work begins: the sampling design, the response design, and the analysis (Congalton and Green, 2009; Stehman and Czaplewski, 1998). These three components are directly influenced by the objectives of the map, i.e., the reasons for its creation. C-CAP data are intended for use at county, regional, and national scales. The data can be used to analyze land cover distribution at single points in time as well as changes across five-year intervals.

Sampling Design

Before beginning the accuracy assessment, the C-CAP team decided to perform nine regional accuracy assessments, which could then be combined into a national-level report (Figure 1). The regions were determined by considering boundaries used during the creation of the data sets, general similarity of land cover and change rates within regions, and similar geographic extents. To meet the objectives of this accuracy assessment, a three-stratum approach was used within each region, including (a) current change, (b) near current (2006-2010) and previously mapped change (1996-2001 and 2001-2006) with the geographic "nearness" determined by region as described below, and (c) the remaining area (Figure 2). From past C-CAP data sets, it was known that change within coastal regions occupies a very small portion of the landscape over a five-year period (approximately four percent). To obtain adequate reference locations within mapped change areas, recommendations by Olofsson *et al*. (2014) were followed and Stratum 1 (black) was created, which contained all pixels of mapped change between 2006 and 2010. Reference locations within this stratum could be used to assess change/no-change accuracy, individual change class accuracy, and overall mapping accuracy. The remaining area in the region (no-change mapped between 2006 and 2010) was then to be split between Strata 2 and 3, described below.

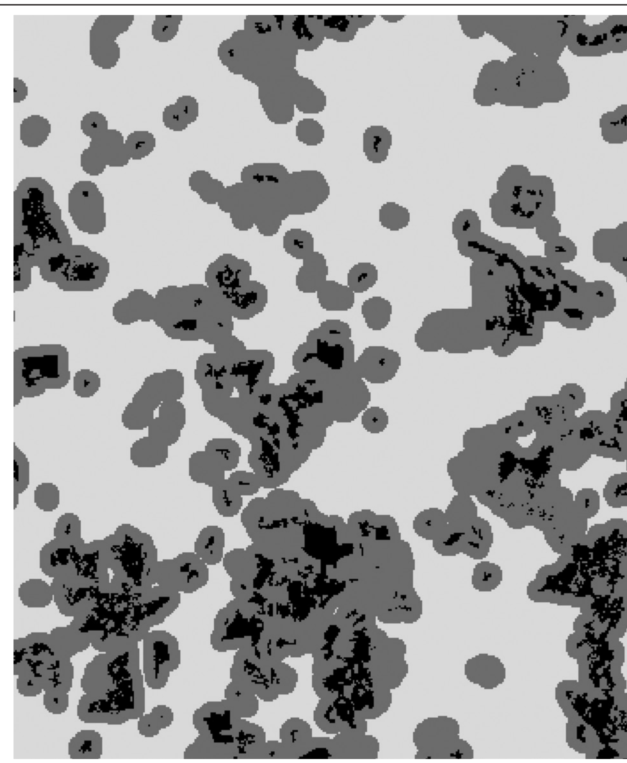

Figure 2. A three- stratum approach was used for placement of reference sites, including current 2006-2010 change (Stratum 1, black), adjacent-to current and past change (1996-2001 and 2001-2006) (Stratum 2, mid-gray), and the remaining area (Stratum 3, light gray background). The buffered distance included in Strata 2 varied per region assessed and ranged from 6 to 19 pixels.

From past C-CAP experience it has been noticed that change is often spatially auto-correlated, which means that new change tends to occur at or near previous change. This can easily be seen in urban expansion (Mertens and Lambin, 2000; Sudhira *et al.,* 2004; Yeh and Li, 2001) or in the clustering of timber activity. Past C-CAP data also have shown the same location changing multiple times. For example, a cultivated field becomes a bare field (before development) and then becomes a neighborhood. Silviculture activities can be easily seen as forests are cut and the stand transitions from grass to shrub and back to forest (Foody *et al.,* 1996; Gregory *et al.,* 1981; Lunetta *et al.,* 2004). This knowledge was used to design Stratum 2 (medium gray in Figure 2) to assess potential missed change as well as overall mapping accuracy. All C-CAP change pixels from 1996-2001, 2001-2006, and 2006-2010 were combined and spatially buffered, until the buffered area plus past change areas approximated the desired target area (half the full assessment area minus 2006-2010 change area). Stratum 1 was then removed from the area to ensure that it was not double-sampled.

The remaining area (not in Strata 1 or 2) became Stratum 3 (dark gray background). Reference locations in this stratum may help identify missed change but would be most useful in assessing wall-to-wall accuracy.

The next step was to determine the quantity, type, and placement of sample units within each stratum. Determining the sample size for this accuracy assessment was a balance of maintaining statistical validity, keeping time and costs in check, and still meeting the objectives. Several authors have discussed methods toward determining appropriate sample size (Congalton, 1988; Hay, 1979; Hord and Brooner, 1976; and van Genderen and Lock, 1978). While these techniques are valid approaches to computing overall map or class accuracy, they were not designed to populate an error matrix. The error matrix provides overall accuracy, individual class accuracies, and the ability to analyze which classes are being confused. To populate an error matrix to analyze mapping accuracy and class confusion, a common "rule-of-thumb" for the number of accuracy assessment locations for land cover mapping is 50 per class (Congalton, 1991; Congalton and Green, 1999). C-CAP has a maximum 25-class land cover scheme (*http://coast.noaa.gov/digitalcoast/_/pdf/ccap_class_scheme.pdf*), although only 23 are present in this mapping effort (there were no Unclassified or Tundra), and not all classes were present in every region (total number of categories ranged from 17 to 22 per region). Occasionally a class may have been present, but was very rare and may not have been a category of overall interest.

Regions were assessed as they were completed, with the western Great Lakes completed first. Within this region, 18 land cover classes were present. Based on the 50 per class target minimum, a total of 900 sample locations was set as the standard per region, with 300 being placed in each stratum. The standard of 900 per region was used for consistency across all regions, even though the total number of classes varied from 17 to 22. This may have resulted in fewer sample units than would be ideal, with the more rare classes (e.g., Aquatic Bed) being under-sampled. Sample units were identified using the ERDAS Imagine® Accuracy Assessment tool with the following criteria: stratified random placement; a minimum of 10 per class (not always met with rare classes); and six out of nine land cover pixels around the location (3 × 3 pixel window) required to be homogenous (or location would be discarded). The targeted number of at least 50 points per class was not always met and is explained in the "Results" section in more detail. The rationale for the homogeneity criteria is explained more in the "Response Design" section.

Response Design

Response design for accuracy assessment refers to the protocols involved when determining agreement between reference and map classifications. Olofsson *et al.* (2014) refers to four major features: spatial unit, the source(s) used for reference classification, labeling protocol for reference classification, and definition of agreement.

To keep the natural heterogeneity of land cover in the landscape, C-CAP land cover does not perform any spatial filtering or aggregating of land cover (i.e., final products are delivered 30 meter pixels). With this in mind, it may seem logical to assess at the single pixel level. At the same time, C-CAP is intended for use at the county, regional, or national scale and should only be used as a screening tool for very local or site-specific management decisions. As a compromise between these two scales (single-pixel mapping unit versus best use of the data at the regional level), it was decided that a 3 × 3 pixel window would be the sample unit. This window size is used to help account for mixed pixels, allow for natural homogeneity in the landscape, and allow for potential registration error among the data sources, and is deemed appropriate based on the scale at which C-CAP data should be used. According to recommendations in Stehman and Wickham (2011), a reference unit containing class heterogeneity may be treated as a homogenous unit if specified thresholds are met. For this assessment, a six out of nine (or greater) pixel land cover agreement within the sample unit was treated as homogenous in the analysis. Users of C-CAP data should be aware of these considerations.

Reference data for the previous C-CAP accuracy assessment (2001 product) were collected through a combination of *in-situ* data collection, reference site collection from airplanes, and photointerpretation of high-resolution imagery (satellite and aerial). This process was labor intensive, expensive, and took considerable time to complete. Reference data sources for the 2010 accuracy assessment were freely available imagery and ancillary data sets, including Landsat, National Agriculture Imagery Program (NAIP), Google Earth™, Bing Maps, National Wetlands Inventory (NWI), and Soil Survey Geographic database (SSURGO). One big advantage to using Landsat, NAIP, and Google Earth was the availability of multiple dates, which assisted in determining land cover change. Another advantage of many of these data sets was their availability as web based services (NAIP, NWI, SSURGO) or stand-along products (Google Earth, Bing Maps), which removed the need to have the data stored locally. The NWI and SSURGO data sets were viewed within Google Earth, allowing interpreters quick and easy access to specific NWI classes and SSURGO hydric rating, drainage class, and flooding frequency, which assisted in identifying wetland classes. The final interpreter's land cover call was based on the assessment of all available data sets to arrive at the best decision, with no ancillary data set being used in an absolute manner.

The 900 reference units per region were randomized and split into three groups of 600 each for interpretation. This splitting of reference units was performed such that each reference unit was interpreted by two independent reviewers. Each reviewer had access to all available reference data mentioned above and was responsible for labeling each reference unit with a primary 2010 land cover call, a fuzzy (alternate) 2010 land cover call (if needed), and a 2006-2010 land cover change/no-change call. Fuzzy calls were only used if the interpreter could not positively identify a single dominant land cover (e.g., natural speckling of land cover classes), or when similar classes could not be positively separated (e.g., shrub is distinguished from forest by a height criteria, which may be difficult to determine from available data sets). The use of fuzzy calls is well documented in the literature when considering land cover mapping (Gopal and Woodcock, 1994; Muller *et al.,* 1998; Wickham *et al.,* 2010; Wickham *et al.,* 2013). After interpretation, if reviewers disagreed on the reference, fuzzy, or change calls, those points were not dropped, but were reviewed and discussed to determine final reference calls.

The 2010 land cover and 2006-2010 binary land cover change values were extracted for the nine pixels that comprised each sample unit location. To be deemed correct, six of the nine sample-unit land cover pixels had to match the reference call (or fuzzy call). The six out of nine criteria was subjectively arrived at as a balance between the desire to allow for natural speckling in the land cover and to not overly bias the accuracy assessment through the use of only homogenous areas. The same criterion was used for change/no-change calls.

Analysis

Error matrices (Congalton *et al.,* 1983) were used to calculate overall accuracy, user's and producer's accuracy, and kappa statistics for the overall map. This first matrix was used for analyzing the accuracy of the wall-to-wall map based on all 8,100 sample units. A second matrix was used to assess the accuracy within *mapped* change areas (Stratum 1), which allows for statements to be made in regard to newly mapped areas. A final matrix using reference sites deemed change by the interpreters was created to determine how well true change was being mapped. Change/no-change was also examined, including overall, user's, and producer's accuracy. These statistics were initially computed for each of the nine regions for creating custom C-CAP accuracy reports (*http://coast.noaa.gov/digitalcoast/publications/regional-reports*) and then combined for this national-level analysis.

Results and Discussion

The distribution of all accuracy assessment sample units, by category and strata, is listed in Table 1. A total of 8,100 sample units were used across the three strata and 23 total land cover classes. Scrub/Shrub received the most accuracy assessment sample units (788), and Estuarine Forested Wetland received the fewest (21). Only three classes contained fewer than the recommended 50 sample units (Congalton, 1991; Congalton and Green, 1999). As might be expected, these three classes also make up the least amount of area mapped within the C-CAP area (approximately one-tenth of one percent). The last two columns in Table 1 can be compared to assess if a class was sampled proportionally to its area. For example, Mixed Forest received 4.7 percent of the accuracy assessment sample units and made up 5.8 percent of the national coastal region. The largest discrepancy is with Evergreen Forest receiving 7.7 percent of the accuracy assessment sample units and making up 14.4 percent of the national coastal region.

The differences between reference units and mapped area (last two columns of Table 1) can generally be tied to the distribution of sample units across the three strata. For example, 7.0 percent of the sample units occurred in mapped Grassland/Herbaceous, while only 4.9 percent of the nation was mapped as this class. Grassland/Herbaceous appears to be much more heavily sampled in Stratum 1 (314 sample units) compared to the other strata, even though Stratum 1 was only 3.7 percent of the C-CAP mapped area. This occurred because Grassland/Herbaceous was one of the most common classes occurring in change areas, thus receiving more sample units in comparison to the other strata where Grassland/Herbaceous consisted of a much smaller percentage.

2010 Land Cover

The target overall accuracy for C-CAP is 85 percent with single class accuracy of 80 percent. The national overall accuracy was 84.0 percent with a kappa value of 0.83. The majority of classes met the C-CAP target specification of 80 percent per class accuracy. Of the ten instances where accuracy was below the targeted 80 percent, all exceeded 70 percent. No classes fell below the 80 percent threshold for both producer and user accuracy. As mentioned earlier, nine regional accuracy assessments were performed to compute statistics per region,

Table 1. Distribution of Accuracy Assessment Sample Units and Land Cover. The Middle Four Columns Show the Distribution of Accuracy Assessment (AA) Sample Units across the Three Strata. The Two Rightmost Columns Display the Overall Distribution of the AA Sample Units and the Overall Distribution of Land Cover. The Last Two Rows Show the Land Area and Percent of the Full Mapping Area.

	Accuracy Assessment Sample Units				Percent of	
Land Cover	Stratum 1	Stratum 2	Stratum 3	Total	Sample Units	Region
Developed, High Intensity (HID)	112	100	90	302	3.7%	0.6%
Developed, Medium Intensity (MID)	150	115	95	360	4.4%	1.4%
Developed, Low Intensity (LID)	158	129	121	408	5.0%	3.1%
Developed, Open Space (OSD)	162	110	105	377	4.7%	1.6%
Cultivated Crops (CULT)	120	185	243	548	6.8%	10.5%
Pasture/Hay (PAST)	96	155	167	418	5.2%	7.0%
Grassland/Herbaceous (GRS)	314	136	116	566	7.0%	4.9%
Deciduous Forest (DEC)	109	218	230	557	6.9%	11.8%
Evergreen Forest (EVR)	183	243	201	627	7.7%	14.4%
Mixed Forest (MIX)	72	169	138	379	4.7%	5.8%
Scrub/Shrub (SS)	437	202	149	788	9.7%	10.7%
Palustrine Forested Wetland (PFW)	91	187	187	465	5.7%	8.8%
Palustrine Scrub/Shrub Wetland (PSS)	141	128	109	378	4.7%	2.6%
Palustrine Emergent Wetland (PEM)	146	124	106	376	4.6%	2.2%
Estuarine Forested Wetland (EFW)		10	11	21	0.3%	0.1%
Estuarine Scrub/Shrub Wetland (ESS)	2	16	21	39	0.5%	0.0%
Estuarine Emergent Wetland (EEM)	33	76	81	190	2.3%	0.9%
Unconsolidated Shore (UCS)	68	80	67	215	2.7%	0.2%
Bare Land (BARE)	153	98	91	342	4.2%	0.7%
Open Water (OW)	116	115	283	514	6.3%	12.4%
Palustrine Aquatic Bed (PAB)	17	37	27	81	1.0%	0.0%
Estuarine Aquatic Bed (EAB)	20	50	51	121	1.5%	0.1%
Perennial Snow (SNOW)		17	11	28	0.3%	0.0%
Total	2700	2700	2700	8100		
Area (square miles)	29,740	387,164	376,559	793,463		
Percent of Region	3.7%	48.8%	47.5%			

then combined for the full report. The overall accuracy and kappa values for each region and the national level are shown in Figure 1. The full error matrix for the nation is shown in Table 2. At the regional level, overall accuracies ranged from a low of 82.3 percent in the Southeast, while the highest was in the western Great Lakes at 85.6 percent.

Major sources of classification confusion as revealed in the error matrix include the following.

Low Vegetation

Pasture/Hay, Cultivated, Grassland, and Open Space Developed were all mapped with some confusion. The confusion between Cultivated and Pasture/Hay is fairly common and has been seen in other maps (Fuller *et al.,* 1994; Vogelmann *et al.,* 2001; Wickham *et al.,* 2010; Wickham *et al.,* 2013). These classes are often best classified through the use of multiple dates of imagery to help detect spectral trends throughout the growing season. Typically, two dates of imagery were available for the 2010 classification, but they were not selected with Cultivated classification as the primary driver and thus may not have been the best available for these classes. Additionally, Pasture and Cultivated are often discriminated by land use, which may be difficult to distinguish through remote sensing alone, especially at the scale used by C-CAP.

Scrub/Shrub, Grassland, and Forest

Timber activity, which often results in the cycling of grassland to scrub/shrub to forest over time, occurs heavily in several regions of the coastal area. Within each region, and the full nation, changes within these classes were the dominant type of change. The scrub/shrub class is generally a transitional class between Grassland and forest classes and is distinguished in C-CAP by a height criterion (woody vegetation less than 5-meter height is scrub/shrub). Depending on the region and species of forest, the transitions between these classes may occur at different rates. It was not uncommon to identify areas of grass to forest change in the Southeast over this five-year time span, where southern yellow pine plantations are common. In the Northeast or Great Lakes, previously cut forest areas may remain in the grass or scrub/shrub category for a much longer time. Many studies have reported that longer gaps in time between mapping efforts result in forest change detection errors (Lunetta *et al.,* 2004; Muchoney and Haack, 1994; Sader *et al.,* 2003). Two studies in particular (Lunetta *et al.,* 2004; Wilson and Sader, 2002) recommended using imagery collected less than three years apart to detect forest harvesting operations and other changes that do not remove the full canopy. Since height cannot be directly measured in the Landsat data, other criteria must be used (tone, texture, shadow, etc.), resulting in the confused classes.

Levels of Development

The developed classes were separated from each other through the application of thresholds to a percent impervious surface (e.g., if the percent impervious was 80 percent or greater, the class was High Intensity Developed). Errors in the percent impervious surface value could translate to errors in the development class label. The majority of the error among these classes was isolated to the immediate neighboring intensity class (e.g., Low Intensity Developed being confused with Open Space Developed or Medium Intensity Developed). Because the percent impervious surface is a spectrally derived value, it is susceptible to variation from spectral differences naturally caused by the time of the year or the reflectivity of different impervious surfaces (e.g., blacktop versus concrete). Thus, two surfaces that should have the same imperviousness percentage, and Development category, may receive different values. More information on the creation of the percent impervious surface and its accuracy can be found in Wickham *et al.* (2010), Wickham *et al.* (2013), and Xian and Homer (2010).

Table 2. Error Matrix for C-CAP 2010 Land Cover. User's and Producer's Accuracy Below Target Values of 80 Percent are in Bold Along the Edges of the Table.

Map	HID	MID	LID	OSD	CULT	PAST	GRS	DEC	EVR	MIX	SS	PFW	PSS	PEM	EFW	ESS	EEM	UCS	BAR	OW	PAB	EAB	SNOW	Total	User
HID	289	2	1	1			2												7					302	95.7%
MID	28	297	11	6	1		5				2			1					9					360	82.5%
LID	1	13	325	22	5	6	9	3	4	1	2	2	1	1			1		8	4				408	79.7%
OSD		1	1	310	2	7	22	5	1	4	11	5	1	2					5					377	82.2%
CULT		1	2	16	455	22	21	3	3	1	12	2	1	5					4					548	83.0%
PAST		1	4	18	43	305	25	4	1	1	13			3										418	73.0%
GRS	1	1	2	9	21	13	455	12	9	5	25	1	1	5					5	1				566	80.4%
DEC			1	1	1	3	4	493	12	17	17	4	1	3										557	88.5%
EVR				2				3	577	17	21	3	2						1				1	627	92.0%
MIX						1	2	14	25	320	15	2												379	84.4%
SS			1	3	3	10	33	44	49	39	578	8	9	4					7					788	73.4%
PFW			2	1	4		2		5	1	2	426	14	5						3				465	91.6%
PSS				1	2	1	3	2	1	2	9	37	300	17			1	1		1				378	79.4%
PEM					4	4	6		1		1	8	15	313		1	4	4	6	6	3			376	83.2%
EFW															20		1							21	95.2%
ESS												1	1	2		28	7							39	71.8%
EEM					1		1	1				2	1	9			171	1		3				190	90.0%
UCS									1					3				167	4	39		1		215	77.7%
BARE		1		6	9	7	12	2	6		7	1		2				1	284	4				342	83.0%
OW			1		1	1		1		1	1	2		3			1		1	500		1		514	97.3%
PAB					1							1		1			2	3		11	59	3		81	72.8%
EAB										1	1									10	1	108		121	89.3%
SNOW											1								1				26	28	92.9%
Total	319	317	351	396	553	380	602	587	695	410	718	505	347	379	20	29	188	177	342	582	63	113	27		
Producer	90.6%	93.7%	92.6%	78.3%	82.3%	80.3%	75.6%	84.0%	83.0%	78.0%	80.5%	84.4%	86.5%	82.6%	100.0%	96.6%	91.0%	94.4%	83.0%	85.9%	93.7%	95.6%	96.3%		84.0%

Bare Land

Bare Land was occasionally classified as developed categories. This is most common when a site is being prepared for development, but construction has not yet begun. The proximity to existing development, and the bright reflectance of the bare soil, creates confusion with developed land (Alberti *et al.,* 2004; Platt and Goetz, 2004; Treitz *et al.,* 1992).

Water and Unconsolidated Shore

Open Water was incorrectly mapped as Unconsolidated Shore. In coastal locations, nearshore wave action, water turbidity, and tidal stage all influence the separation of these two classes. Examination of these incorrect sample locations shows that the Unconsolidated Shore class is most likely overmapped in general, often because of wave action and nearshore sediment present in the imagery.

The use of fuzzy or alternate calls was allowed in conditions where the field class was either difficult to positively identify (e.g., Cultivated versus Pasture, shrub versus 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. Fuller *et al.* (1994) demonstrated that by removing class boundary pixels (edge pixels) from their Landsat classification, their final accuracy rose from 46 percent to 71 percent. Analysis showed that 36.2 percent (2,932) of the sample units received a fuzzy call, and of the 6,806 correctly mapped sample units, 1,120 (16.5 percent) were correct due to the fuzzy call. A matrix plotting "primary reference call" versus "fuzzy reference call" was constructed to analyze the specific nature of reference class confusion. As was expected, the majority of the fuzzy calls occurred within the four developed classes, among the low vegetation classes (Grassland, Cultivated, and Pasture), between shrub and Grassland, and between shrub and forested.

2006-2010 Change

To date there is no standard approach to assessing categorical land cover change maps. One of the more common approaches is to use the standard confusion matrix with the land cover change categories representing the row and column values (Congalton and Green, 2009; Khorram, 1999). The 2006 NLCD used a modified version of this approach to assess 2001-2006 land cover change (Wickham *et al.,* 2013). For their study, Multi-Resolution Land Characteristics Consortium partners selected a subset of 22 land cover change categories of interest, which were then assessed. However, for C-CAP, a full change assessment would possibly represent a 625 class × 625 class matrix, which is not practical or economical. Instead of attempting to identify priority change categories, it was decided to assess change/no-change accuracy using a 2 × 2 matrix (Congalton and Green, 2009; Morisette and Khorram, 2000), with the hopes that the stratification method used would allow statements to be made about mapped or missed change.

Overall, C-CAP's change/no-change accuracy was 88.7 percent (Table 3). Committed change (falsely identified change) was the largest error with a user accuracy of 71.2 percent (777 sample locations mapped as change, but deemed no-change by the reviewers), which is much higher than omitted change (93.3 percent producers accuracy). The 777 committed error locations were assessed in their own error matrix and resulted in 74.6 percent overall accuracy. This seems to indicate that the method used to identify potential change pixels (i.e., creating the change mask) may be overestimating change, but the methods used to assign a land cover class are reasonably accurate. C-CAP is willing to accept higher rates of committed change error (opposed to omitted change error) as long as these locations are accurately mapped, and thus improve the overall product. These locations and trends of committed change may be used in future editing efforts, since they are indicative of potential errors with the 2006 map.

Two additional land cover error matrices were created to compare different categories of change based on Table 3. The overall statistics are shown in Table 4. The error matrix for reference units in mapped change areas (Table 5) allowed for statements about the accuracy of the current mapping effort and indicated an 82.3 percent accuracy. This value is slightly lower than the overall map accuracy. The second category of change was those sample units where both the map and the interpreter identified change (1,923 sample units). The accuracy of these was 86.3 percent. Land cover confusion within these matrices was similar to the overall map discussed previously. It may be useful to compare these two values to the accuracy of the committed sample units (74.6 percent) in the preceding paragraph. The 2006 land cover was used during the classification of the 2010 map as training data, and errors within that map (as indicated by the 777 sample units) may affect the training data and any trend modeling that may take place (e.g., 2006 shrub transitioning to 2010 forest). The net effect was that incorrect 2006 land cover data reduced the 2010 land cover accuracy either due to lesser quality training data or the inability to accurately model logical land cover transitions.

Assessing mapped change is a fairly straightforward task, but assessing missed change is problematic. Of the 5,400 total sample units in mapped no-change areas, only 137 were deemed missed change. Stratum 2 (specifically designed to try to identify potential missed change) contained 107 (78.1 percent) of these missed change locations. After conducting the change analysis, the team felt that overall change had been somewhat overcalled, although there were limited missed true change sites as well. As designed, Stratum 2 appeared to be helpful in identifying potentially missed change in C-CAP land cover. Because C-CAP land cover updates are created by classifying potential change areas only, and keeping non-change areas as currently mapped, there is a preference to slightly over-map potential change at the beginning, so as to not miss change features. As the potential change features are mapped, ideally areas of no-change will fall out because the newer land cover call will match the older land cover call, resulting in no change.

Table 3. Error Matrix for Change and No-Change for the 2006-2010 C-CAP Land Cover Change Product.

		Reference			
		No Change	Change	Total	Users
Map	No Change	5263	137	5400	97.5%
Map	Change	777	1923	2700	71.2%
	Total	6040	2060	8100	
	Producers	87.1%	93.3%		88.7%

Table 4. Overall Land Cover Accuracy from Different Change Categories. These Sample Units were Subset from Locations of Mapped Change and Can be Used as an Indicator of the Land Cover Accuracy of Current Mapping Processes.

		From Error Matrix	
Change Category	Sample Units	Correct	Overall Accuracy
All Mapped Change	2700	2223	82.3%
Mapped Change = Reference Change	1923	1659	86.3%

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 percent overall and 80 percent per class. Nine regional accuracy assessments were performed on the 2010 C-CAP data. These nine reports were combined to produce this national-level accuracy report. The combined accuracy for the nation was 84.0 percent, with the majority of individual classes exceeding 80 percent accuracy. There were no classes with accuracy below 80 percent for both user and producer accuracy.

A unique sampling strata approach was used in an effort to assess mapped change areas as well as make some statements regarding potential missed change. Change/no-change accuracy for the nation was 88.7 percent, with commission errors being the largest component. It was found that 74.6 percent of the false change locations received the correct 2010 call, indicating that the new classification approaches appear to be working fairly well.

Although the accuracy did not meet the targeted 85 percent, the overall quality of the map was high. During the 2010 land cover update cycle, the C-CAP team expended considerable effort to improve the mapping accuracy and consistency of development and wetland classes across the nation. As C-CAP completes the next round of land cover updates, improvements to other land cover classes will be incorporated as deemed appropriate. Each regional accuracy report highlights several of these areas for improvement.

References

Alberti, M., R. Weeks, and S. Coe, 2004. Urban land-cover change analysis in central Puget Sound, *Photogrammetric Engineering & Remote Sensing*, 70(9):1043–1052.

Civco, D., J. Hurd, E. Hoffhine Wilson, C. Arnold, and M. Prisloe, 2002. Quantifying and describing urbanizing landscapes in the Northeast United States, *Photogrammetric Engineering & Remote Sensing*, 68(10):1083–1090.

Cleckner, H., and T.R. Allen, 2014. Dasymetric mapping and spatial modeling of mosquito vector exposure, Chesapeake, Virginia, USA, *ISPRS International Journal of Geo-Information*, 3(3):891–913.

Congalton, R., 1988. Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data, *Photogrammetric Engineering & Remote Sensing*, 54(5):587–592.

Congalton, R., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37(1):35–46.

Congalton, R., and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, First edition, Lewis Publishers, Boca Raton, Florida, 137 p.

Congalton, R., and K. Green, 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Second edition, Taylor and Francis, Boca Raton, Florida, 183 p.

Congalton, R., R. Oderwald, and R. Mead, 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques, *Photogrammetric Engineering & Remote Sensing,* 49(12):1671–1678.

Dobson, J., E. Bright, R. Ferguson, D. Field, L. Wood, K. Haddad, H. Iredale, J. Jensen, V. Klemas, R. Orth., and J. Thomas, 1985. *NOAA Coastal Change Analysis Program (C-CAP): Guidance for Regional Implementation*, NOAA Technical Report NMFS (National Marine Fisheries Service) 123, U.S. Department of Commerce, 92 p.

Foody, G.M., G. Palubinskas, R.M. Lucas, P.J. Curran, and M. Honzak, 1996. Identifying terrestrial carbon sinks: Classification of successional stages in regenerating tropical forest from Landsat TM data, *Remote Sensing of Environment*, 55(3): 205–216.

Fuller, R.M., G.B. Groom, and A.R. Jones, 1994. The land cover map of Great Britain: An automated classification of Landsat Thematic Mapper data, *Photogrammetric Engineering & Remote Sensing*, 60(5):553–562.

Table 5. Error Matrix Using Only Sample Units Occurring in Mapped Change Areas.

Map \ Reference	HID	MID	LID	OSD	CULT	PAST	GRS	DEC	EVR	MIX	SS	PFW	PSS	PEM	EEM	ESS	UCS	BAR	OW	PAB	EAB	Total	User
HID	106	1		1			2											2				112	94.6%
MID	14	110	5	5	1		4				1			1				9				150	73.3%
LID	1	3	106	14	4	2	8	2	1	1	1	2	1	1				8	3			158	67.1%
OSD				131	2	3	12	3		1	2	2		2				4				162	80.9%
CULT				2	90	9	9				5	1		2				2				120	75.0%
PAST				2	7	73	11	1			2											96	76.0%
GRS	1			1	3	3	282	3	2	1	11	1		2				4				314	89.8%
DEC			1		1		2	85	5	2	10	1		2								109	78.0%
EVR									168	1	12		2									183	91.8%
MIX								1	14	47	10											72	65.3%
SS					2	2	22	11	19	9	356	2	6	2				6				437	81.5%
PFW			1		1		2		3		1	77	4	1					1			91	84.6%
PSS						1	2				2	10	115	9		1	1					141	81.6%
PEW						2	4		1			1	3	123		1	2	5	2	2		146	84.2%
ESS															2							2	100.0%
EEW													1	5		26	1					33	78.8%
UCS														1			59		7		1	68	86.8%
BAR		1		1	2		5		1		2	1					1	139				153	90.8%
OW					1	1		1			1	1		2				1	108			116	93.1%
PAB																				16	1	17	94.1%
EAB																					20	20	100.0%
Total	122	115	113	157	114	96	365	107	214	62	416	99	132	153	2	28	64	180	121	18	22	2700	
Producer	86.9%	95.7%	93.8%	83.4%	78.9%	76.0%	77.3%	79.4%	78.5%	75.8%	85.6%	77.8%	87.1%	80.4%	100%	93%	92.2%	77.2%	89.3%	88.9%	90.9%		

Correct	2223
Percent Correct	82.3%

Galbraith, H., R. Jones, R. Park, J. Clough, S. Herrod-Julius, B. Harrington, and G. Page, 2002. Global climate change and sea level rise: Potential losses of intertidal habitat for shorebirds, *Waterbirds*, 25:173–183.

Greene, C.M., K. Blackhart, J. Nohner, A. Candelmo, and D.M. Nelson, 2014. A national assessment of stressors to estuarine fish habitats in the contiguous USA, *Estuaries and Coasts*, 38(3):782–799.

Gregory, M.S., S.J. Walsh, and J. Vitek, 1981. Mechanics of monitoring forest clearcuts and their regeneration, *Proceedings of the LARS Symposia*, Paper 461, URL: *http://docs.lib.purdue.edu/lars_symp/461* (last date accessed: 28 June 2016).

Hansen, M.C., and T.R. Loveland, 2012. A review of large area monitoring of land cover change using Landsat data, *Remote Sensing of Environment*, 122:66–74.

Hay, A.M., 1979. Sampling designs to test land use map accuracy, *Photogrammetric Engineering & Remote Sensing,* 45(40):529–533.

Hord, R.M., and W. Brooner, 1976. Land use map accuracy criteria, *Photogrammetric Engineering & Remote Sensing,* 42(5):671–677.

Kang, M., V. Mesev, and S. Myint, 2014. Urbanization and quality of stormwater runoff: Remote sensing measurements of land use in an arid city, *Korean Journal of Remote Sensing*, 30(3):399–415.

Khorram, S. (editor), 1999. *Accuracy Assessment of Remote Sensing-Derived Change Detection*, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, 64 p.

Liang, L., and P. Gong, 2015. Evaluation of global land cover maps for cropland area estimation in the conterminous United States, *International Journal of Digital Earth*, 8(2):102–117.

Lowe, M., and M. Peterson, 2014. Effects of coastal urbanization on salt-marsh faunal assemblages in the northern Gulf of Mexico, *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 6(1):89–107.

Lu, D., and Q. Weng, 2006. Use of impervious surface in urban land-use classification, *Remote Sensing of Environment*, 102(1):146–160.

Lunetta, R.S., D.M. Johnson, J.G. Lyon, and J. Crotwell, 2004. Impacts of imagery temporal frequency on land-cover change detection monitoring, *Remote Sensing of Environment*, 89(4):444–454.

Margriter, S., G. Bruland, G. Kudray, and C. Lepczyk, 2014. Using indicators of land-use development intensity to assess the condition of coastal wetlands in Hawai'i, *Landscape Ecology*, 29:517–528.

Mertens, B. and E.F. Lambin, 2000. Land-cover-change-trajectories in southern Cameroon, *Annals of the Association of American Geographers*, 90:467–494.

Millette, T., E. Marcano, and D. Laflower, 2014. Winter distribution of moose at landscape scale in northeastern Vermont: A GIS analysis, *Alces*, 50:17–26.

Morisette, J.T., and S. Khorram, 2000. Accuracy assessment curves for satellite-based change detection, *Photogrammetric Engineering & Remote Sensing*, 66(7):875–880.

Morris, J.T., P.V. Sundareshwar, C.T. Nietch, B. Kjerfve, and D.R. Cahoon, 2002. Responses of coastal wetlands to rising sea level, *Ecology*, 83(10):2869–2877.

Muchoney, D.M., and B.N. Haack, 1994. Change detection for monitoring forest defoliation, *Photogrammetric Engineering & Remote Sensing*, 60:1243–1251.

Nagendra, H., D.K. Munroe, and J. Southworth, 2004. From pattern to process: Landscape fragmentation and the analysis of land use/land cover change, *Agriculture, Ecosystems & Environment*, 101(2):111–115.

National Oceanic and Atmospheric Administration (NOAA), Economics: National Ocean Watch (ENOW). Data based on data from the Bureau of Labor Statistics and the Bureau of Economic Analysis, Charleston, South Carolina: NOAA Office for Coastal Management, URL: *http://www.coast.noaa.gov/digitalcoast/data/enow/* (last date accessed: 28 June 2016).

National Oceanic and Atmospheric Administration (NOAA), Coastal Change Analysis Program (C-CAP). Regional Land Cover Database, data collected 1995-present, Charleston, South Carolina: National Oceanic and Atmospheric Administration (NOAA) Office for Coastal Management, URL: *http://www.coast.noaa.gov/digitalcoast/data/ccapregional/* (last date accessed: 28 June 2016).

Nestlerode, J.A., V.D. Hansen, A. Teague, and M.C. Harwell, 2014. Application of a three-tier framework to assess ecological condition of Gulf of Mexico coastal wetlands, *Environmental Monitoring and Assessment*, 186(6):3477–3493.

Olofsson, P., G.M. Foody, M. Herold, S.V. Stehman, C.E. Woodcock, and M.A. Wulder, 2014. Good practices for estimating area and assessing accuracy of land change, *Remote Sensing of Environment*, 148:42–57.

Platt, R.V., and A.F. Goetz, 2004. A comparison of AVIRIS and Landsat for land use classification at the urban fringe, *Photogrammetric Engineering & Remote Sensing*, 70(7):813–819.

Porter, J.H., R.D. Dueser, and N.D. Moncrief, 2015. Cost-distance analysis of mesopredators as a tool for avian habitat restoration on a naturally fragmented landscape, *The Journal of Wildlife Management*, 79(2):220–234.

Raghavan, R., K. Almes, D. Goodin, J. Harrington, Jr, and P. Stackhouse, Jr, 2014. Spatially heterogeneous land cover/land use and climatic risk factors of tick-borne feline *cytauxzoonosis, Vector-Borne and Zoonotic Diseases*, 14(7):486–495.

Sader, S.A., M. Bertrand, and E.H. Wilson, 2003. Satellite change detection of forest harvest patterns on an industrial forest landscape, *Forest Science*, 49(3):341–353.

Stehman, S.V., and R.L. Czaplewski, 1998. Design and analysis for thematic map accuracy assessment: Fundamental principles, *Remote Sensing of Environment*, 64(3):331–344.

Stehman, S.V., and J.D. Wickham, 2011. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy, *Remote Sensing of Environment*, 115:3044–3055.

Sudhira, H.S., T.V. Ramachandra, and K.S. Jagadish, 2004. Urban sprawl: Metrics, dynamics and modelling using GIS, *International Journal of Applied Earth Observation and Geoinformation*, 5(1):29–39.

Treitz, P.M., P.J. Howarth, and P. Gong, 1992. Application of satellite and GIS technologies for land-cover and land-use mapping at the rural-urban fringe: A case study, *Photogrammetric Engineering & Remote Sensing*, 58(4):439–448.

U.S. Census Bureau. 2011. Census 2010, URL: *http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml* (last date accessed: 28 June 2016).

van Genderen, J.L., B.F. Lock, and P.A. Vass, 1978. Remote sensing: Statistical testing of thematic map accuracy, *Remote Sensing of Environment*, 7(1):3–14.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel, 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogrammetric Engineering & Remote Sensing*, 67(6):650–662.

Wickham, J.D., S.V. Stehman, J.A. Fry, J.H. Smith, and C.G. Homer, 2010. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States, *Remote Sensing of Environment*, 114(6):1286–1296.

Wickham, J.D., S.V. Stehman, L. Gass, J. Dewitz, J.A. Fry, and T.G. Wade, 2013. Accuracy assessment of NLCD 2006 land cover and impervious surface, *Remote Sensing of Environment*, 130:294–304.

Wilson, E.H., and S.A. Sader, 2002. Detection of forest harvest type using multiple dates of Landsat TM imagery, *Remote Sensing of Environment*, 80:385–396.

Woods and Poole Economics, Inc., 2011. Complete economic and demographic dataset, URL: *http://www.woodsandpoole.com/main.php?cat=country*, data processed by NOAA to determine coastal county summary totals and absolute and percent change (last date accessed: 28 June 2016).

Xian, G., and C. Homer, 2010. Updating the 2001 National Land Cover Database impervious surface products to 2006 using Landsat imagery change detection methods, *Remote Sensing of Environment*, 114(8):1676–1686.

Yeh, A.G.O., and X. Li, 2001. Measurement and monitoring of urban sprawl in a rapidly growing region using entropy, *Photogrammetric Engineering & Remote Sensing*, 67(1):83–90.

(Received 13 January 2016; accepted 08 March 2016; final version 29 April 2016)