Oregon and Washington 2010 Coastal Change Analysis Program Accuracy Assessment

National Oceanic and Atmospheric Administration (NOAA)
Coastal Services Center

May 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

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 Oregon and Washington, covering nearly 74,000 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, new land cover classes have been added, 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 Oregon and Washington C-CAP land cover update was conducted through the contract vehicle at the NOAA Coastal Services Center. The 2010 land cover classification 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 January 2014.

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

- The overall accuracy for the Oregon and Washington 2010 C-CAP product was 84.9% (0.84 kappa).
- Two classes fell below 80% for both producer[1] *and* user[2] accuracy; five classes were below 80% for producer accuracy, and eight were below 80% for user accuracy (Table 2).
- The accuracy for change/no-change was 88.1%, with the largest error being committed change (70.3% accuracy). It is interesting to note that of these committed change locations (falsely mapped as change) the accuracy was 73.0% for the 2010 call, indicating the 2006 call was incorrect.
- Of the 300 sample locations in mapped change areas, the accuracy was 75.3%.

Methods

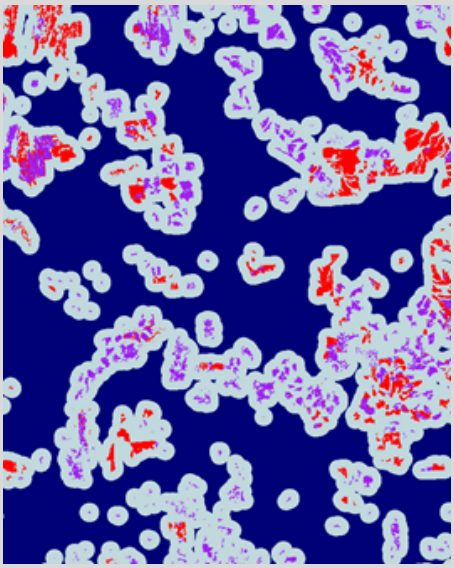
The C-CAP team met and discussed accuracy assessment on multiple occasions and determined three essential requirements:

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 6 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 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 would be 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, Evergreen Forest received the most accuracy assessment sample units (129) and Palustrine Aquatic Bed received the fewest (18). 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 5.1% of the accuracy assessment (AA) sample units and comprised 5.8% of the region. The largest discrepancy is with Evergreen Forest receiving 14.3% of the AA sample units and comprising 47.6% of the region. If pure stratified random sampling were used, this difference would have been smaller, but because a minimum of 10 points per class-per stratum was chosen, sample locations were allocated to the remaining classes, instead of being concentrated in Evergreen Forest.

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	13	10	10	33	3.7%	0.3%
Developed, Medium Intensity	13	11	11	35	3.9%	0.9%
Developed, Low Intensity	15	12	12	39	4.3%	2.1%
Developed, Open Space	15	11	12	38	4.2%	0.6%
Cultivated Crops	13	11	12	36	4.0%	1.7%
Pasture/Hay	13	13	14	40	4.4%	4.1%
Grassland/Herbaceous	43	15	15	73	8.1%	6.5%
Deciduous Forest	13	12	11	36	4.0%	1.8%
Evergreen Forest	20	55	54	129	14.3%	47.6%
Mixed Forest	15	18	13	46	5.1%	5.8%
Scrub/Shrub	41	27	18	86	9.6%	14.9%
Palustrine Forested Wetland	5	11	11	27	3.0%	1.0%
Palustrine Scrub/Shrub Wetland	13	11	10	34	3.8%	0.5%
Palustrine Emergent Wetland	15	11	10	36	4.0%	0.6%
Estuarine Emergent Wetland	0	10	10	20	2.2%	0.1%
Unconsolidated Shore	10	10	11	31	3.4%	0.5%
Bare Land	18	11	12	41	4.6%	1.6%
Open Water	15	11	25	51	5.7%	8.9%
Palustrine Aquatic Bed	0	10	8	18	2.0%	0.0%
Estuarine Aquatic Bed	10	10	10	30	3.3%	0.1%
Perennial Snow	0	10	11	21	2.3%	0.4%
Total	300	300	300	900		
Area (square miles)	2,830	35,440	35,601	73,871		
Percent of Region	3.8%	48.0%	48.2%			

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 84.9% (0.84 kappa). The majority of classes met the C-CAP target specification of 80% per class accuracy. Of the 13 instances where accuracy was below the targeted 80%, 10 of these did exceed 70%. Two classes, Cultivated and Pasture/Hay fell below the 80% threshold for both producers and user accuracy. These two classes were mostly confused with each other. Other common classes with confusion

were Scrub/Shrub being mapped as upland forest categories and Grassland. Grassland was confused with Cultivated, Pasture/Hay, and lower levels of development. Bare Land was also confused with developed classes. The class with the lowest single accuracy was Palustrine Aquatic Bed (55.6% user accuracy). This value may be questionable because of the low number of reference locations (Table 2).

Although most classes did not have more than 50 sample units (the coarse "rule-of-thumb" for accuracy assessment), four classes did exceed 50, and seven classes were over 40. Palustrine Aquatic Bed was the only class with fewer than 20 sample locations, which can be explained by the limited area this class represents in the landscape.

Other major sources of classification confusion were seen in the error matrix:

1. **Low vegetation** – Pasture/Hay, Cultivated, Grassland, and Scrub/Shrub were all mapped with some confusion. The confusion between Cultivated and Pasture/Hay is fairly common and has been seen in other C-CAP regions. 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.
2. **Scrub/Shrub, Grassland, and Upland Forest** – Oregon and Washington have a considerable amount of timber activity, which results in the cycling of Grassland to Scrub/Shrub to Forest over time. 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.
3. **Low levels of development** – New development and urban expansion often falls into Low Intensity and Open Space Developed. Confusion was seen between these two classes, but also with Pasture/Hay and Grassland. The confusion with Pasture/Hay and Grassland can be expected since these two land cover classes are frequently the land converted to development.
4. **Bare Land** – True 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.

Table 2. Full error matrix for the 2010 Oregon and Washington 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.

Reference

Map	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	Estuarine Emergent Wetland	Unconsolidated Shore	Bare Land	Open Water	Palustrine Aquatic Bed	Estuarine Aquatic Bed	Perennial Snow	Total	Users
Developed, High Intensity	30																3					33	90.9%
Developed, Medium Intensity	4	27					2										2					35	77.1%
Developed, Low Intensity	1	3	24	4		2	1		1								3					39	61.5%
Developed, Open Space				32		1	2				3											38	84.2%
Cultivated Crops				1	26	3	3				2						1					36	72.2%
Pasture/Hay				2	6	29	3															40	72.5%
Grassland/Herbaceous							63		1		5			1			3					73	86.3%
Deciduous Forest			1	1				26	2	2	4											36	72.2%
Evergreen Forest								1	123		4						1					129	95.3%
Mixed Forest									6	35	5											46	76.1%
Scrub/Shrub						2	5		6	3	69						1					86	80.2%
Palustrine Forested Wetland									1		1	22	2	1								27	81.5%
Palustrine Scrub/Shrub Wetland					1	1							27	3	1	1						34	79.4%
Palustrine Emergent Wetland														33	1		2					36	91.7%
Estuarine Emergent Wetland															20							20	100.0%
Unconsolidated Shore									1							29		1				31	93.5%
Bare Land							1				1						39					41	95.1%
Open Water																		51				51	100.0%
Palustrine Aquatic Bed														1	2	3		1	10	1		18	55.6%
Estuarine Aquatic Bed																				30		30	100.0%
Perennial Snow											1						1				19	21	90.5%
Total	35	30	25	40	33	38	80	27	141	40	95	22	29	39	24	33	56	53	10	31	19	900	
Producers	85.7%	90.0%	96.0%	80.0%	78.8%	76.3%	78.8%	96.3%	87.2%	87.5%	72.6%	100.0%	93.1%	84.6%	83.3%	87.9%	69.6%	96.2%	100.0%	96.8%	100.0%		84.9%

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 27% of the sample units received a fuzzy call, these calls were rarely responsible (12%) for a location being deemed mapped as correct.

Table 3. Fuzzy reference calls for the 2010 Oregon and Washington C-CAP region.

Fuzzy Reference Calls		
Of the 900 sample locations, 241 (26.8%) had a fuzzy call	For the 764 correctly mapped locations, 93 (12.2%) were correct based on the fuzzy land cover call (671 were correct based on primary call)	Land cover classes with most fuzzy calls include different levels of development; Scrub/Shrub and Grass; upland forest categories; Unconsolidated Shore and Estuarine Aquatic Bed

2006-2010 Change

Overall change/no-change accuracy was 88% (Table 4). Committed change was the largest error with a user accuracy of 70% (89 sample locations mapped as change, but deemed no-change by the reviewers). These 89 locations were assessed in their own error matrix and resulted in 73.0% 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 18 were deemed missed change. Sixteen 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, although there were limited missed true change sites as well.

Table 4. Change/no-change matrix for the 2010 Oregon and Washington 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	582	18	600	97%
	1	89	211	300	70%
	Total	671	229	900	
	Producers	87%	92%		88%

A final analysis was performed using only sample locations interpreted as change (229 locations). Table 5 shows that the overall accuracy of these locations was 78.2%, slightly lower 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

Direct comparisons with the 2001 accuracy assessment cannot realistically be made, since there are different land cover classes and classification methods varied considerably. The accuracy assessment locations (2,208 sites) for the 2001 data set were drawn from a mix of field sample locations along roadways and interpreted from aerial and satellite imagery (Figure 2). The overall accuracy for the 2001 map was 86.1%. There are a couple reasons for the apparent decrease in accuracy from 2001 to 2010. The 2010 map includes more detailed developed classes and the Pasture/Hay class, all of which are difficult to map accurately. Also, the reference sites for 2010 were randomly located, as opposed to the 2001 placement method along roads or clustering within image footprints. Sampling along roadways 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. Finally, accuracy assessment locations for the 2001 map were restricted to 3 x 3 homogenous pixel windows, which typically will remove variable (or “speckled”) landscapes.

Table 5. Error matrix for the 2010 Oregon and Washington 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	Unconsolidated Shore	Bare Land	Open Water	Total
Developed, High Intensity	8															1		9
Developed, Medium Intensity		7					2											9
Developed, Low Intensity	1	1	4				1									3		10
Developed, Open Space				1														1
Cultivated Crops				1	8		1				1							11
Pasture/Hay					1	4	2											7
Grassland/Herbaceous							35		1		4			1		2		43
Deciduous Forest			1					6	1	1	4							13
Evergreen Forest									15		2							17
Mixed Forest									1	5	3							9
Scrub/Shrub							5		3		35					1		44
Palustrine Forested Wetland											1	3						4
Palustrine Scrub/Shrub Wetland													7		1			8
Palustrine Emergent Wetland														10		2		12
Unconsolidated Shore															4			4
Bare Land							1									17		18
Open Water																	10	10
Total	9	8	5	2	9	4	47	6	21	6	50	3	7	11	5	26	10	229

Correct	179
Percent Correct	78.2%

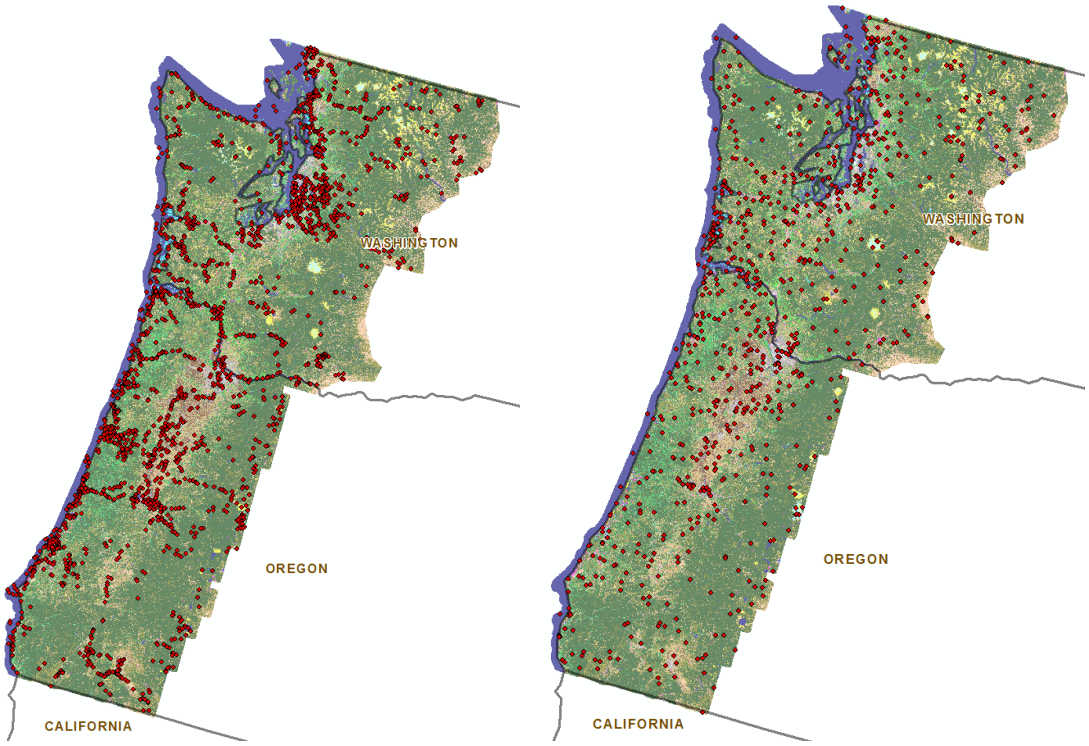


Figure 2. Differences in accuracy assessment site selection method and number of locations can be seen by comparing the 2001 C-CAP Oregon and Washington accuracy assessment locations (left) against the 2010 C-CAP Oregon and Washington accuracy assessment locations (right). Sampling along road networks (linear features) in the 2001 assessment can be seen within the left map versus the stratified random sample for the 2010 approach.

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 Oregon and Washington 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 84.9%, with the majority of individual classes exceeding 80% accuracy.

There were only two classes, Cultivated and Pasture/Hay, with accuracy below 80% for both user and producer accuracy. Change/no-change accuracy for the product was 88.1%, with committed change being the largest error. It was found that 73% of the false change locations received the correct 2010 call, indicating the classification approaches appear to be working fairly well.

Although the accuracy did not meet the target 85%, the overall quality of the map was high. Potential improvements could be performed on the map in the future, or as part of the next update cycle. Confusion among Grassland, Scrub/Shrub, and Forest remains a difficult issue to address. This is especially common due to the amount of timber activity in the region. 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.

Confusion among Pasture/Hay, Grassland, and Cultivated remains a difficult issue to address. 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. These will continue to be difficult class breaks to make, and more discussions should be held to address this issue.

The final issue that could be addressed is confusion of Bare and developed classes. An impervious surface product, provided by the U.S. Geological Survey (USGS), served as the base for the developed classes. USGS performed considerable edits to improve this product, including the removal of false development, during the 2010 mapping effort. These improvements were limited to east of the Mississippi River, since time and cost did not allow for further editing. Significant improvements were seen where this editing occurred, and repeating this process on the West Coast C-CAP data sets may result in similar results.