Comparison of NLCD with NWI Classifications of Baldwin and Mobile Counties, Alabama

By Larry Handley and Chris Wells

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2009

Contents

Figures

Tables

Conversion Factors

SI to Inch/Pound

Multiply	By	To obtain
Length		
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
meter (m)	1.094	yard (yd)
Area		
hectare (ha)	2.471	acre
square meter (m^2)	10.76	square foot (ft^2)

Comparison of NLCD with NWI Classifications of Baldwin and Mobile Counties, Alabama

By Larry Handley and Chris Wells

Introduction

Mobile Bay National Estuary Program (MBNEP) requested an assessment of the accuracy of National Land Cover Data 2001 (NLCD 2001, compiled by the Multi-Resolution Land Characteristics Consortium) as compared to National Wetlands Inventory (NWI [Cowardin and others, 1979], to include uplands through incorporation of modified Anderson Level III definitions [Anderson and others, 1976]) mapping of Mobile and Baldwin Counties conducted by the U.S. Geological Survey's (USGS) National Wetlands Research Center (NWRC). Both classifications were checked against stratified randomly selected ground-based sites and with each other to compare the accuracy of the NLCD with NWI classification methods. Previous to this comparison, we conducted an accuracy assessment of the modified NWI in Louisiana with an overall accuracy of 93 percent (Handley and others, 1995).

For this accuracy comparison, numerous visits were made by photointerpreters to the Mobile Bay area to establish signatures for the modified NWI mapping. After all of the quadrangles in Baldwin and Mobile Counties were mapped, an accuracy assessment of those maps was conducted by field-checking the mapped classes with ground-based check sites. These same sites were used to check the accuracy of the NLCD and also as the basis for cross-walking the NLCD and NWI classification systems for direct comparison of the two methods.

This accuracy assessment focused on the latest of several dates of habitat mapping projects of the Mobile Bay area by NWRC using the same photointerpretation methodology and classification system. We have used this methodology and NWI classification on numerous other dates and regions. Extrapolation beyond this data set is appropriate for maps of the same methodology and classification of the Mobile Bay region. This accuracy assessment provides analytical information about the classification accuracy of 2001/2002 NWI classification and the 2001 NLCD classification.

Accuracy of NWI by Ground-Based Check Sites

A purely random sampling of the area was not practical because of time and legal constraints on the USGS, which must have written permission to enter private property. Thus, all ground-based check sites (hereafter referred to as "check sites") had to be publicly accessible and timely. For these reasons, all check sites were located adjacent to public roadways.

All roads in a geographic information system (GIS) layer of Baldwin and Mobile Counties greater than 5 km in length were stratified by county and within counties, by north and south of Interstate 10, resulting in four sampling areas. This sampling scheme facilitated wide distribution of sampled roads between counties and north (into the hills) and south (into the coastal flatland). Thus, the area was roughly subdivided for sampling into four quadrants. The interstate was chosen for rapid accessibility across both counties. The geographic area is so large that time and funding played a significant role in both ground-truthing and accuracy assessment.

Selection of check sites within quadrants was accomplished by randomly ordering the roads greater than 5 km in length and then sequentially reading down the list. Removing shorter roads from consideration decreased the probability of entering urban areas, which have a disproportionately large number of short roads which would have biased the sample toward urban areas. From the beginning, a stop was made every 0.3 mi (measured by using the vehicle's odometer in tenth-mile intervals, 0.1 mi = 0.26 km), and the opposite side of the road from the previous stop was evaluated. The side of the road for analysis at the beginning was randomly selected.

Once at the sample point, a Global Positioning System (GPS) position was recorded for the position 100 m perpendicular to the road by using a laser distance-measuring device connected to the GPS receiver (fig. 1). One observer classified this point. A second observer paced 100 m parallel to the road going forward, and a third observer paced 100 m parallel to the road going back. The second and third observers recorded their classifications of the land at their respective points. All three observers then evaluated the center position, came to a classification agreement, and recorded it on paper. Data for the map position, its polygon designation, all three of the observers' classes, and the final consensual classification were recorded on a laptop. After the sample was completed, the observers drove another interval and repeated the procedure until the end of the road was reached.
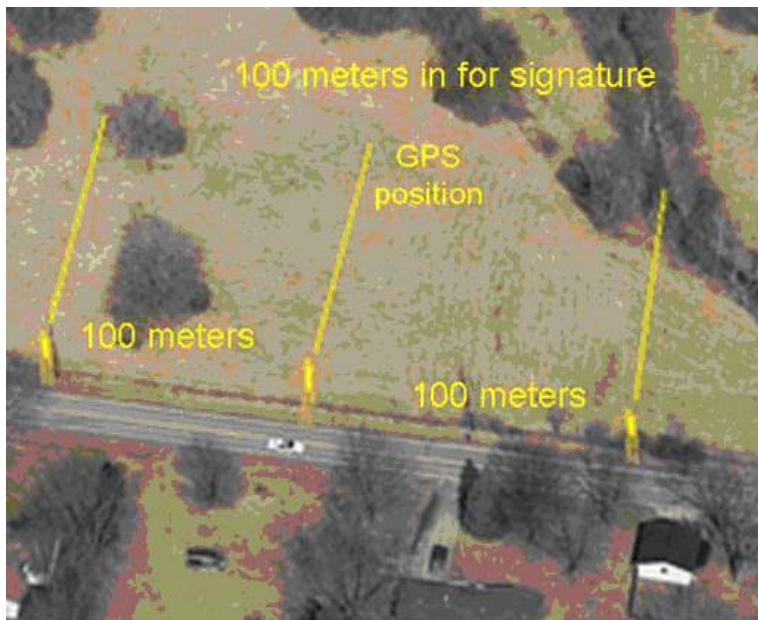

Figure 1.   Representation of sampling method. See description in text for methodology.

Every evening in the field, the day's observations were uploaded to the statistician at NWRC. The statistician determined whether the 95 percent confidence interval had been met for each level of the NWI hierarchy. From this information it could be determined how many days of observations in the field were required to attain the sampling confidence interval at each level of the classification hierarchy.

Analysis

After the requisite number of samples for the 95 percent confidence interval had been attained, accuracy was determined by comparing the NWI polygon definition of the map with the consensual ground sample. Accuracy was the total number correct of a given class divided by the total number in that class.

The hierarchy was defined thus, from top to bottom:

1. Upland versus wetland
2. System
3. Class
4. Subclass
5. Modifier

A difficulty with the method was that the NWI photography was flown during the 2000–2002 time period but that the ground sampling began in 2004 and was completed in 2008. In that period of time, land was cleared for new houses and new roads; grasslands converted to farmland or scrub-shrub and scrub-shrub to forest. We addressed this complication by inferring what the area looked like 3–7 years earlier, as well as what we would call it at the date of ground observation. In the majority of instances the check sites appeared unchanged, and so, the delineation stood on its own. In a few cases a forest may have been clear-cut and replaced by a subdivision or other feature.

Under these circumstances, a choice was made whether to ignore the check site, record it as incorrect, or second-guess the determination. The decision was made by consensus. Most often the site was skipped and was not included in the accuracy assessment. It was not recorded as incorrect unless it was certain that the original delineation itself was incorrect before the habitat change. On occasions when the original delineation was clearly correct it was recorded as such.

Results of NWI as Compared to Check Sites

A total of 195 accuracy assessment field checks were conducted, each with three independent observations and one consensual observation. The observations were roughly half in each county, and about half of each county was sampled in the north and about half in the south to ensure wide coverage of the samples.

The class "upland versus wetland" consisted of only those two possibilities. By definition, all areas designated as urban or agriculture were uplands. The accuracy was evaluated at each level of the classification hierarchy. With the uppermost level, the analysis was one-way—we looked only at the errors of calling a habitat an upland when it was actually a wetland to determine whether we were missing wetlands. For the class "upland versus wetland," accuracy was 98 percent; standard error was 1 percent.

"System" included all NWI system categories. There were no upland designations for "system." Accuracy was 97 percent; standard error was 1 percent.

"Class" consisted of 15 NWI and upland categories. Accuracy was 88 percent; standard error was 3 percent.

"Subclass" consisted of eight NWI and upland categories. Accuracy was 88 percent; standard error was 3 percent.

"Modifier" is only found in the NWI. These designations resolve differences among seasonality, duration of saturation, and tidal frequency. Accuracy was 78 percent; standard error was 6 percent.

Cross-Walking NWI and NLCD Classification Schemas

Definitions of each category within each classification schema were compared. The NLCD categories have definitions that match various levels of NWI. Because of its complexity, NWI required collapsing lower levels of the hierarchy until the categories were sufficiently general to match the other schema. Even then the correspondence was not necessarily one to one, as some definitions defied cross-categorization. The resultant classification cross-walk is provided in table 1.

Table 1. Habitat description and cross-walked classification codes for the National Land Cover Data (NLCD) and National Wetlands Inventory (NWI) classification systems.

Description	NLCD code	NWI code (modified)
Open water	11	Most E1, L1, P, R1, R2 with AB and UB)
Developed, open space	21	UUp
Developed, all densities	23	UU except UUp
Barren land	31	UB (all)
Unconsolidated shore	32	E2US, L2US, MUS, R1US, R2US
Deciduous forest	41	UF6 (all)
Evergreen forest	42	UF7 (all)
Mixed forest	43	UF8 (all)
Scrub/shrub	52	USS (all)
Grassland/herbaceous	71	UR (all)
Cultivated crops/pasture	82	UA
Palustrine forest	91	PF (all)
Palustrine scrub/shrub	92	PSS (all)
Estuarine scrub/shrub	94	E2SS(all)
Palustrine emergent (persistent)	96	PEM (all)
Estuarine emergent	97	E2EM (all)

The NLCD 2001 map was transformed to Universal Transverse Mercator (UTM) projection coordinates to match the NWI maps. The NLCD file was converted from a raster (*.img) file to a GRID file and then to a polygon (shape) file by using ArcMap (ESRI, Inc., Redlands, Calif.) GIS. During conversion, the original cell boundaries were maintained so that the derived polygon (shape) file perimeters matched the boundaries of the original raster file. In both files an additional column was added for the cross-walk to the NLCD. The NLCD classes had to be reduced somewhat in upland categories because our modified NWI does not have as many categories for uplands as does NLCD. Also,

there were some issues with the open-water categories that were resolved by collapsing the water categories of both NWI and NLCD.

Comparison of NLCD with Check Sites

We also compared the sites sampled for the NWI accuracy assessment (described above) to the NLCD classification. Of the 16 NLCD classes cross-walked, 14 had check sites associated with them that cross-walked to NLCD. For two of the NLCD classes, "estuarine scrub/shrub" and "developed open space," there were no associated check sites. NLCD did not classify any "estuarine scrub/shrub" (code 94) or "unconsolidated shore" (code 32) in Mobile Bay.

The NWI classes of the check sites were cross-walked to the NLCD classification by using the method in table 1. There was 38 percent agreement between the NLCD classes and the cross-walked NWI check site classifications. Agreement came in "open water" (9 percent), "developed, all densities" (48 percent), "evergreen forest" (41 percent), "mixed forest" (15 percent), "scrub/shrub" (36 percent), "cultivated crops/pasture" (70 percent), and "palustrine forest" (46 percent). Check site accuracy is provided in table 2.

Table 2. Accuracy of National Land Cover Data (NLCD) map to check sites. See text for description of check site sampling methodology.

Description	Code	Check sites per class	Check site accuracy
Open water	11	11	9%
Developed, all densities	23	31	48%
Barren land	31	1	0%
Deciduous forest	41	4	0%
Evergreen forest	42	45	41%
Mixed forest	43	13	15%
Scrub/shrub	52	11	36%
Grassland/herbaceous	71	8	0%
Cultivated crops/pasture	82	10	70%
Palustrine forest	91	48	46%
Palustrine emergent (persistent)	96	10	0%
Estuarine emergent	97	2	0%

The average accuracy, excluding NLCD classes not sampled, was 27 percent. The small sample size in some NLCD classes may have increased the likelihood of sampling error having a greater effect. Unlike the NWI design for ground visits, no test was done to evaluate the effect of check site sampling size on NLCD classification. Consequently, no confidence limits are provided.

NLCD Accuracy as Compared to NWI

Two comparisons were made between NWI and NLCD. The first was a strict polygon labeling comparison, and the second was polygon labeling weighted by area.

Polygon Labeling Accuracy

NWI codes are hierarchical and have the potential for hundreds of combinations to describe a given polygon in Mobile Bay. There were a total of about 87,000 NWI polygons and 382 unique classifications in the two-county area. By comparison, NLCD (16 classes in Mobile Bay and 24,200,000 cells) is a far simpler classification system but has about 300 times the number of polygons. We simplified NWI by aggregating 382 categories upward to match the 16 NLCD classes by using the cross-walk described in table 1. Adjacent NWI polygons with identical labels were combined into single polygons. Similarly, identical labels in adjacent NLCD polygons were combined into single polygons.

A matrix detailing the origin and disposition of the NLCD area within the NWI polygons is provided in table 3. Obviously, there is strong agreement between NWI and the NLCD concerning "open water" (class 11). There is low to no agreement between NWI and NLCD in the category "unconsolidated shore" (class 32); in NLCD most of that category went to "open water."

Table 3.   Matrix indicating the class tested across the top (the column headers are the particular National Land Cover Data [NLCD] classes within the classification) and the cross-walked National Wetlands Inventory (NWI) category down the left-hand column (again, the labels indicate the NLCD classes).

[Units are in hectares. Ideally, all of the area should be found along the diagonal (top left to bottom right). The diagonals are indicated in boldface type. The values off-diagonal indicate the area (in hectares) misclassified as some habitat type other than the co-occurrence diagonal. The accuracy column (labeled "correct") of the cross-walked classifications was derived by dividing the diagonal values (bold) by the row sums]

NLCD vs. modified NWI

	11	21	23	31	32	41	42	43	52	71	82	91	92	94	96	97	correct
11	**212976**	583	460	1409	0	7	491	51	324	187	388	2144	0	0	2781	0	96%
21	0	**404**	86	5	0	0	6	1	3	20	65	3	0	0	1	0	68%
23	189	25465	**21266**	1255	0	76	5265	490	6938	2785	14970	3093	0	0	606	0	26%
31	113	217	338	**880**	0	1	67	10	190	271	638	86	0	0	87	0	30%
32	616	45	31	872	**0**	0	47	4	23	29	63	116	0	0	507	0	0%
41	37	299	27	7	0	**103**	360	377	533	84	329	1723	0	0	32	0	3%
42	267	10064	937	194	0	1165	**183276**	20305	26899	3494	4991	45875	0	0	1101	0	61%
43	26	1775	176	41	0	334	4387	**2081**	2544	390	1641	8847	0	0	117	0	9%
52	15	1319	283	101	0	370	8905	1394	**18927**	6355	2544	3111	0	0	186	0	44%
71	95	3697	1267	608	0	94	3195	452	9446	**5023**	11869	2305	0	0	406	0	13%
82	23	4336	797	93	0	24	1012	97	4064	3755	**66992**	1167	0	0	171	0	81%
91	1507	1167	185	57	0	288	28393	3092	1874	589	1186	**85840**	0	0	1825	0	68%
92	45	114	27	6	0	16	1910	140	269	100	89	5481	**0**	0	421	0	0%
94	10	13	10	5	0	0	14	0	3	3	0	254	0	**0**	508	0	0%
96	85	172	138	75	0	14	590	57	554	455	537	1791	0	0	**522**	0	10%
97	297	123	56	139	0	0	72	0	65	55	11	440	0	0	9318	**0**	0%

The summary results of polygon labeling of a one-tailed comparison of NWI with NLCD are given in table 4.

Table 4. Summary of National Land Cover Data versus National Wetlands Inventory in Mobile Bay. Excluding the absent classes (see text for explanation), accuracy increased to 43 percent.

[stderr, standard error]

mean	stderr	
0.3186981	0.08551	
0.2914461	0.34595	95% confidence interval
		(95% of the values will lie between 0.2914 and 0.34595)

There is about 32 percent agreement with 95 percent of the values falling within the interval 0.29–0.35. Perfect agreement would have a value of one. A brief scan of the "correct" column in table 3 indicates that some categories had very high agreement (NLCD versus NWI in category 11, "open water," has a value of 0.96, very high agreement). At the other end, classes 32, 92, 94, and 97 had none correct. This result is due to the NLCD of Mobile Bay not classifying "unconsolidated shore," "wetland scrub/shrub," or "estuarine emergent." The classification comparison in table 3 calculates the mean of the correct values column. Those four absent classes greatly bias the error. Removal of those columns from the calculation raises the agreement to about 43 percent (95 percent confidence interval was 0.38 to 0.47).

The fundamental difference between NWI photography and interpretation compared to NLCD satellite imagery classification methodology resulted in maps with strikingly different lineaments of the boundaries even when they fundamentally agreed on a classification. The NLCD boundaries zigzagged, while the NWI boundaries were smoothly curved. Thus, line overlap results in abundant sliver polygons as the NWI line overlays the zigzagged NLCD line. There are various methods for smoothing NLCD, but they result in compounding errors during interpolation.

The decision was made to leave the boundaries as they were in the original maps and deal with the boundary alignment during analysis. In this first analysis, all polygons received equal weight. This method resulted in sliver polygons having weight equal to large polygons. A review of polygon areas revealed that almost half of the polygons were less than 0.1 ha, which is about the size of a 30- by 30-m pixel in NLCD. Thus, about half of the pixels were slivers, probably produced as an artifact of the map type difference: raster versus vector. The pixels were retained for the polygon labeling accuracy assessment. Sliver polygons were created most frequently from polygons with complicated boundaries, especially those with large perimeter to area ratios (common with rivers and streamside wetlands).

We completed another analysis in which polygon area weighted the results (table 5). In this case, polygons with larger areas had a proportionately larger influence on the comparison than did smaller polygons. By accounting for area, the overall accuracy doubled to about 63 percent. Though the analysis of the individual contributions of the habitats is incomplete, we expect that they will indicate increased accuracy for the

polygons that typically have large perimeter to area ratios, as the slivers from wetland areas lose influence in the analysis.

Table 5. Accuracy of National Land Cover Data (NLCD) as compared to National Wetlands Inventory (NWI) maps of Baldwin and Mobile Counties, Ala. This analysis was weighted by the area of each polygon.

Observations	Accuracy	Variance	Standard deviation
1	0.62786	0.23365	0.48338

Discussion

The Multi-Resolution Land Characteristics Consortium (MRLC) planned for an NLCD accuracy assessment of regional classifications for the 2001 classification (Vogelmann and others, 1998; see also *http://www.epa.gov/mrlc/accuracy-2001.html*), but according to the MRLC Web site this has not been conducted. Thus, there is no objective, independent accuracy comparison of the 2001 NLCD. It should be remembered that NLCD is intended as an analysis of national and regional coverage for any given date and as input for time-series changes over large areas. The classification system of NLCD was developed for small-scale classification aggregated by general cover classes that are reasonably accurate at those scales.

This accuracy assessment is dependent on a methodology and analysis developed for the unique characteristics and properties of NWI classification, which is mapped to polygon, as compared to NLCD, which uses raster classification and mapping. This cartographic presentation results in fundamentally different classification boundaries. Overlaying these map types cannot prevent boundary differences resulting in a tremendous proliferation of sliver polygons (fig. 2). The intersection of both maps resulted in three-quarters of a million polygons, nearly half of which were less than the 900-m^2 pixel size used in NLCD mapping (the smaller of the minimum mapping units of NWI and NLCD versus NWI mapping was about 0.5 ha). It is the authors' opinion that this cartographic difference needs to be more fully explored before accepting the accuracy difference between NWI and NLCD presented here at face value.
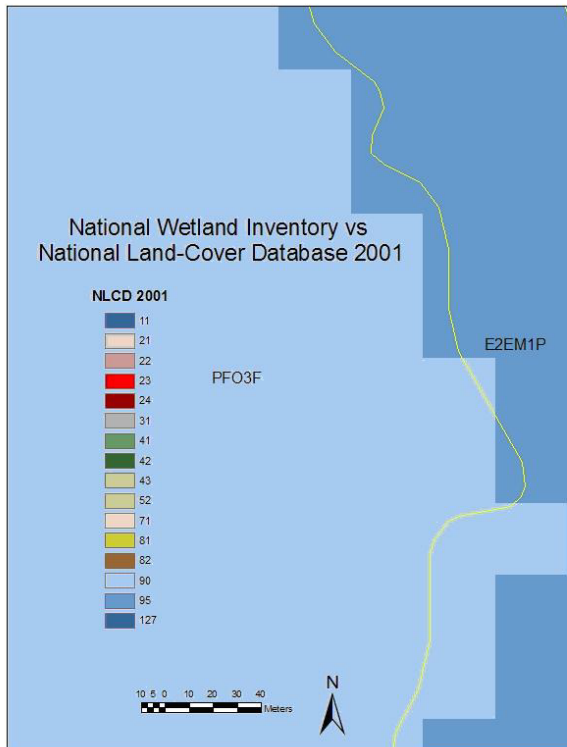
Figure 2. Expanded portion of Mobile Bay indicating the fundamental problem of boundary alignment between National Wetlands Inventory (vector, yellow line) and National Land Cover Data (raster blocks in shades of blue).

The authors are satisfied that the cross-walking matrix was appropriate, given the constraints of the two classification methods. Some differences in classes are simply irreconcilable because they do not exist in one classification system or the other. These missing classes greatly contribute to accuracy differences.

Acknowledgments

References

Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E., 1976, A land use and land cover classification system for use with remote sensor data: U.S. Geological Survey Professional Paper 964.

Cowardin, L.M., Carter, V., Golet, F.C., and LaRoe, E.T., 1979, Classification of wetlands and deepwater habitats of the United States: U.S. Fish and Wildlife Service FWS/OBS-79/31.

Handley, L., Fuhrmann, M.J., Fuller, D.A., and Blackmon, J.H., Jr., 1995, An accuracy assessment—the 1988 Louisiana habitat maps: Charlotte, N.C., American Society of Photogrammetry and Remote Sensing.

Vogelmann, J.E., Sohl, T.L., Campbell, P.V., and Shaw, D.M., 1998, Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources: Environmental Monitoring and Assessment, v. 51, p. 415–428.