Accuracy assessment of the National Land Cover Database Impervious Surface dataset for Roanoke, Virginia

Tammy E. Parece, Ph.D. Candidate & Dr. James B. Campbell, Department of Geography, Virginia Tech

April 4, 2014

Virginia Tech GIS & Remote Sensing 2014 Research Symposium

Abstract

The Multi-Resolution Land Charateristics Consortium (MRLC) developed National Land Cover Database Impervious Surface (NLCD IS) data to identify percent developed imperviousness for the coterminous USA. We present the results of an accuracy assessment on this data for the City of Roanoke, Virginia. First, we performed a classic accuracy assessment using a set of random points generated by GIS, and high resolution aerial photographs (1/2 foot resolution), varying the NLCD IS' percent imperviousness from 10% to 75% per cell, resulting in an overall accuracy of around 70% for most thresholds. Then a polygon impervious surface dataset was delineated in GIS using the same high resolution aerial photos, and subsequently subdivided into 30 meter by 30 meter pixels matching each cell boundary of the NLCD IS data. A second accuracy assessment was performed on a cell by cell basis, comparing the NLCD IS to this newly created impervious surface dataset. Finally, terrain relief, specifically percent slope created from a 30 meter digital elevation model, was added to the analysis to determine if it impacted the accuracy of the NLCD IS data in the cell by cell assessment.

Background and Study Area

Land cover changes involving increasing amounts of impervious surfaces cause alterations of the hydrologic cycle (prevents infiltration and increases runoff into water bodies) (DeBusk et al. 2010; Welker et al. 2010) and affects natural temperature regulation (causing urban areas to be warmer than surrounding rural areas) (Slonecker et al. 2001; Geiger et al. 2003). Mapping impervious surfaces is essential for evaluating these impacts and implementing effective environmental and urban management planning (Slonecker et al. 2001; Civco et al. 2002; Bauer et al 2005) .

The National Land Cover Database Impervious Surfaces (NLCD IS) was derived by means of regression tree software using both leaf-on and leaf-off Landsat images and images from NOAA's Defense Meteorological Satellite Program [Fry et al. 2006]. The dataset presents a continuous layer with a gradient of imperviousness from 0 to 100 percent for each 30 meter by 30 meter pixel (MRLC 2012) (Figure 1). Accuracy assessments have been performed on the full NLCD with resulting overall accuracies ranging from 59.7% to 80.5% (Yang et al. 2001), 43%–83% (Stehman et al. 2003), and 78%–85% (Wickham et al. 2013). We were unable to locate a similar accuracy assessment performed specifically on the NLCD IS dataset.

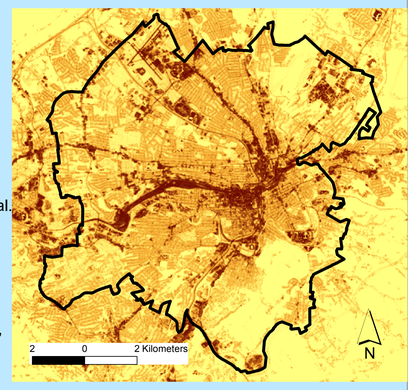


Figure 1. NLCD IS overlaid with boundary of Roanoke, Virginia, yellow = 0% graduating to dark brown for 100% impervious

The City of Roanoke, Virginia is located in a valley between the Blue Ridge Mountains and the Alleghany Highlands (Figure 2). Roanoke, the largest metropolitan region in southwestern Virginia, is characterized by a variety of urban land uses. The city's history is largely based upon its role as a regional transportation hub for rail and road traffic with services and industries supporting the rail system, as well as finance, distribution, trade, manufacturing, and health care businesses.

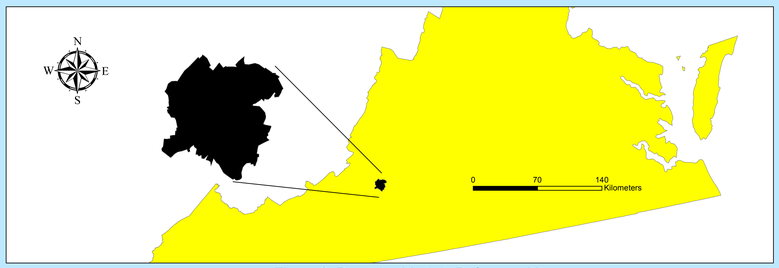


Figure 2. Roanoke, Virginia Reference Map

Methods

Classic Accuracy Assessment:
1) Developed a set of threshold maps for the NLCD IS data set. We set the thresholds at 10%, 25%, 30%, 40%, 45%, 50% and 75%.
2) Created a reference data set by generating random points in ArcGIS and using 2011 high-resolution aerial photos (1/2 foot resolution) from the Virginia Base Mapping Program, identified each point as either impervious or other on the aerial photo and each IS threshold (from step 1).
3) Calculated overall accuracy and kappa for each NLCD IS threshold.

Cell-by-Cell Comparison
1) Manually delineated impervious surfaces from the same high-resoulation aerial photos.
2) Created a fishnet in ArcGIS using the NLCD IS as a reference layer so that the grid cells would match the size and boundaries of the NLCD IS, then used Indentity tool in ArcGIS to calculate percent impervious, as delinated from the aerial photos, for each grid cell.
3) Extracted the percent impervious for each NLCD IS grid cell that matched the exact location of each cell in Step 3.
4) Added percent slope (as calculated from the DEM) to each grid cell.
5) Calculated the difference between the NLCD IS and manually delinated values, and performed a regression analysis (slope = independent variable) to determine slope's influence on this difference.

Results - Classic Accuracy Assessment

Figures 3 - 6 are the results of varying the threshold of the NLCD IS percent impervious from 10% to 75%. The threshold percent represents the minimum value of imperviousness set for each grid cell.

Figure 3. NLCD IS, 10% and above

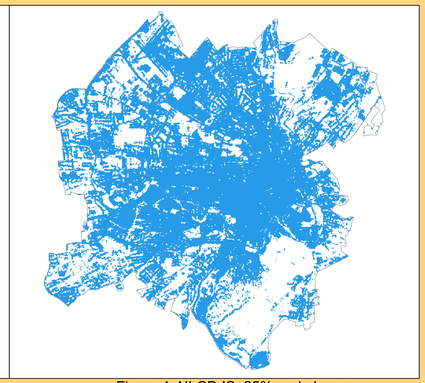


Figure 4. NLCD IS, 25% and above

Figure 5. NLCD IS, 50% and above

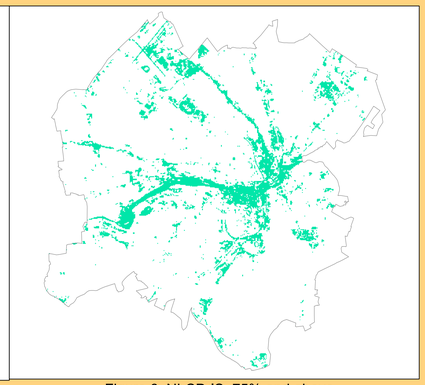


Figure 6. NLCD IS, 75% and above

Table 1 reveals the accuracy assessment results for various NLCD IS thresholds. From threshold values of 35% to 75%, overall accuracy hovers around 70% and kappa 0.40. The highest overall accuracy and kappa was achieved when setting the NLCD IS threshold at 45% imperviousness.

Table 1. Classic accuracy assessment and kappa for different NLCD IS thresholds

Impervious Threshold	Overall Accuracy	kappa
10%	53.5%	0.21
35%	68.6%	0.39
40%	71.4%	0.42
45%	72.8%	0.43
50%	72.6%	0.39
75%	70.6%	0.25

Results - Cell-by-Cell Accuracy Assessment

Figure 7 shows that the fishnet created for a cell size of 30 x 30 meters matches exactly with NLCD IS cell size and boundaries.

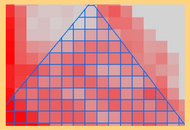


Figure 7. Fishnet grid superimposed on NLCD IS cells

Figure 8 shows impervious surfaces (magenta) for the entire City of Roanoke delineated from high-resolution aerial photos. Delineation has been completed for 75% of the City (black boundary) and partially done for the remaining 25%. The area in black will be used for cell-by-cell comparison.

The number of cells from the NLCD that covers this area totals 93,393. Table 2 shows a sample of the results of extracting NLCD IS, aerial IS, slope percent, and IS difference for each of the 93,393 fishnet grid cells.

Table 2. NLCD IS, aerial IS, slope percent, and IS difference for each cell

OID *	Shape *	CID	NLCD_IS	slope_per	Aerial_IS	IS_Diff
43302	Point	43301	44	21.04089	100	56
43612	Point	43611	38	10.91447	92	54
43609	Point	43608	74	0.843141	73	1
43927	Point	43926	29	15.72552	62	33
44241	Point	44240	55	14.17189	40	15
44243	Point	44242	36	19.60476	16	20
43610	Point	43609	72	14.4699	12	60
44242	Point	44241	59	17.08268	0	59

(8 out of 96130 Selected)

random

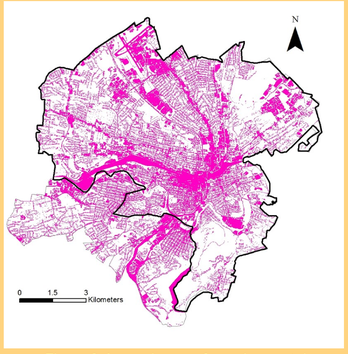


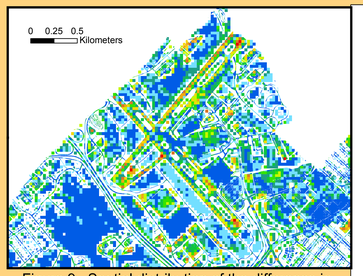
Figure 8. Impervious surfaces (magenta) delineated from aerial photos. Black border is the study area.

Our overall accuracy for all pixel values of NLCD IS as compared to the reference data set from the aerial photos was only 11.8%. The producers' error ranged from 38.0% gradually decreasing to 0% as the IS increased from 0 to 100% IS). Our RMSE was 1384.1.

Since the overall accuracy was so low, we did not calculate kappa, but instead categorized the percent IS and completed a second accuracy assessment. Table 3 shows the results of this additional assessment. The percent IS for both data sets were categorized 0 - 9, 10 - 19...90 - 100. Using a range of values in our assessment increased the overall accuracy to 33.6%, still a rather low value, so again kappa was not calculated. Of particular note, we found that again the producers' error decreased as the percent IS increased for each pixel. Figure 9 shows the spatial distribution of the difference in the per pixel percent IS (aerial minus NLCD IS, larger values = greater disagreement).

Table 3, Accuracy Assessment results when categorizing the percent IS for NLCD IS data and high resolution aerial photo refernce dataset.

NLCD IS	N	Aerial 0 - 9	Aerial 10 -19	Aerial 20 -29	Aerial 30 - 39	Aerial 40 - 49	Aerial 50 - 59	Aerial 60 - 69	Aerial 70 - 79	Aerial 80 - 89	Aerial 90 to 100	
0 to 9	20515	17865	1006	773	417	193	111	61	27	24	38	87.08%
10 to 19	9248	5718	1136	910	534	325	211	119	88	70	137	61.8%
20 to 29	8488	3631	1282	1156	943	604	333	162	129	91	157	42.8%
30 to 39	11890	3236	1853	1935	1702	1328	866	383	205	139	243	27.2%
40 to 49	14625	2121	1702	2166	2391	2389	1960	842	409	254	391	14.5%
50 to 59	9781	917	617	977	1272	1601	1648	1056	587	395	711	9.4%
60 to 69	5251	420	227	294	430	531	657	654	542	459	1037	8.0%
70 to 79	4087	277	139	182	195	253	352	411	423	505	1350	6.8%
80 to 89	4177	178	91	108	144	189	215	327	450	580	1895	4.3%
90 to 100	5331	77	34	43	62	102	128	188	299	538	3860	1.4%
N	93393	34440	8087	8544	8090	7515	6481	4203	3159	3055	9819	36.9%
		51.9%	12.4%	9.0%	5.2%	2.6%	1.7%	1.5%	0.9%	0.8%	0.4%	31413
												33.6%



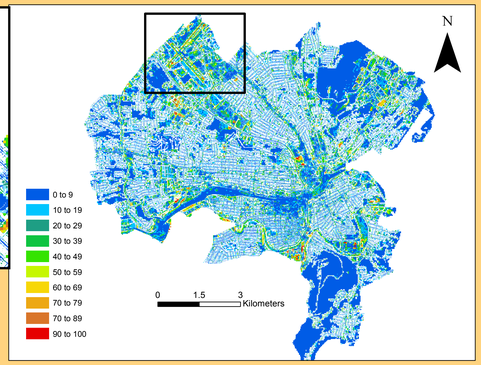


Figure 9. Spatial distribution of the difference in per pixel value (aerial - NLCS IS), (detail shows Roanoke/Blacksburg International Airport). Smaller values (blues) represent agreement, larger values (reds) higher disagreement

Results of the regression analysis on the impact of slope on the difference between the reference data set and the NLCD IS data set produced an R-squared value of less than 1%. A separate regression analysis using percent tree canopy cover produced similar results.

Conclusions

- We anticipated higher accuracies at the lowest and highest levels of IS, but we had decreasing accuracies from low to high percents, shadowing from buildings and topography along with concealment by tree canopy contributed to the inability to discern impervious surfaces.
- Mixed pixels caused much more of a confounding effect than anticipated. The NLCD IS layer is a generalization process, not a discrete identification process as completed in vector delineation.
- Our aerial identification was conducted with a much finer resolution (1/2 foot) than the NLCD IS (30 meter), thus we have better precision and detail.
- Some of the differences between the NLCD IS and the aerial photo interpreted IS can be attributed to differences in data acquisition dates. NLCD IS dataset was completed in 2006 and the aerial interpretation was accomplished using 2011 aerial photos.
- The MRLC advises that NLCD datasets are more accurate at regional and national levels, than at a local level.
- In addition, The MRLC acknowledges that the 2006 NLCD IS dataset is inaccurate and a new dataset with 2011 data is scheduled for release this month (April 2014).

References:

Bauer, M.E.; Loeffelholz, B.; Wilson, B. Estimation, Mapping and Change Analysis of Impervious Surface Area by Landsat Remote Sensing. In Proceedings of Pecora 16—Global Priorities in Land Remote Sensing, Sioux Falls, SD, USA, 22–27 October, 2005.

Civco, D.L.; Hurd, J.D.; Wilson, E.H.; Arnold, C.L.; Prisloe, M.P., Jr. Quantifying and describing urbanizing landscapes in the Northeast United States. Photogramm. Eng. Remote Sens. 2002, 68:1083–1090.

DeBusk, K., W. F. Hunt, U. Hatch and O. Sydorovych (2010). "Watershed retrofit and management evaluation for urban stormwater management systems in North Carolina." Journal of Contemporary Water Research and Education 146: 64-74.

Fry, J.; Xian, G.; Jin, S.; Dewitz, J.; Homer, C.; Yang, L.; Barnes, C.; Herold, N.; Wickham, J. Completion of the 2006 National Land Cover Database for the Conterminous United States. Photogramm. Eng. Remote Sens. 2011, 77: 858–864.

Geiger, R., R. H. Aron and P. Todhunter (2003). The Climate Near the Ground, 6th Edition. Lanham, Md., Rowman & Littlefield.

MRLC (Multi-Resolution Land Charateristics Consortium National Land Cover Database). Available online: http://www.mrlc.gov/nlcd06_data.php (accessed on 19 December 2012).

Slonecker, E. T., D. B. Jennings and D. Garofalo (2001). "Remote Sensing of impervious surfaces: A review." Remote Sensing Reviews 20(3): 227-255.

Stehman, S.V.; Wickham, J.D.; Smith, J.H.; Yang, L. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results. Remote Sens. Environ. 2003, 86: 500–516.

Welker, A. L., B. M. Wadzuk and R. G. Traver (2010). "Integration of education, scholarship, and service through stormwater management." Journal of Contemporary Water Research and Education 146: 83-91.

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

Yang, L.; Stehman, S.V.; Smith, J.H.; Wickham, J.D. Thematic accuracy of MRLC land cover for the eastern United States. Remote Sens. Environ. 2001, 76: 418–422.

The NLCD IS data set was downloaded from the MRLC website; the City of Roanoke boundary file was downloaded from the City of Roanoke, Virginia FTP site; 2011 High resolution aerial photos were accessed from the Virginia Geographic Information Network GIS server.