Removing rural roads from the National Land Cover Database to create improved urban maps for the United States, 1992-2011

By Christopher E. Soulard, William Acevedo, and Stephen V. Stehman

Published in: Photogrammetric Engineering and Remote Sensing, vol. 84, issue 2, February 2018, p. 101-109

Full Title

Removing rural roads from the National Land Cover Database to create improved urban maps for the United States, 1992-2011

Description

Rural roads were removed from the National Land Cover Database developed classes and wall-to-wall maps for the conterminous United States were created for four years (1992, 2001, 2006, and 2011) to better characterize urban development and urban change.

Abstract

Quantifying change in urban land provides important information to create empirical models examining the effects of human land use. Maps of developed land from the National Land Cover Database (NLCD) of the conterminous United States include rural roads in the developed land class and therefore overestimate the amount of urban land. To better map the urban class and understand how urban lands change over time, we removed rural roads and small patches of rural development from the NLCD developed class and created four wall-to-wall maps (1992, 2001, 2006, and 2011) of urban land. Removing rural roads from the NLCD developed class involved a multi-step filtering process, data fusion using geospatial road and developed land data, and manual editing. Reference data classified as urban or not urban from a stratified random sample was used to assess the accuracy of the 2001 and 2006 urban and NLCD maps. The newly created urban maps had higher overall accuracy (98.7%) than the NLCD maps (96.2%). More importantly, the urban maps resulted in lower commission error of the urban class (23% versus 57% for the NLCD in 2006) with the trade-off of slightly inflated omission error (20% for the urban map, 16% for NLCD in 2006). The removal of approximately 230,000 km^2 of rural roads from the NLCD developed class resulted in maps that better characterize the urban footprint.

These urban maps are more suited to modeling applications and policy decisions that rely on quantitative and spatially explicit information regarding urban lands.

Keywords: developed; urban, roads; land use; land cover; land change; NLCD; United States

1. Introduction

Urban land is defined in a variety of ways, but is generally distinguished from other forms of developed land based on higher population density, higher building density, higher land use intensity, and/or more impervious cover. For the purposes of this article, urban is defined as densely developed clusters of land including residential, commercial, and industrial land uses. We define “rural roads” as all roads located outside of urban areas, as well as small patches of rural development such as mixed-use agricultural communities, low-density outposts, mining drill pads, railroads, and highway rest stops. We subsume all non-urban land types under the label “rural roads” to simplify terminology because rural roads represent the vast majority of area represented by these types of non-urban development. Monitoring the status, trends, and spatial patterns of urban lands is essential to understanding the causes and consequences of human land use practices on the landscape. By linking urban changes to spatiotemporal factors and policy drivers (Bengston et al., 2004; Wang et al., 2012), scientists, planners, and land managers may be able to anticipate future changes on the landscape. Similarly, understanding how urban changes directly impact wildlife habitat (including habitat loss and fragmentation) (Fahrig, 2003; Swenson and Franklin, 2000), surface and ground-water hydrology (Göbel et al., 2004; Strauch et al., 2008), and surface albedo (Taha et al., 1988), as well as how changes indirectly alter species biodiversity and air and water pollution (Fahrig, 2003; Strauch et al., 2008) may help land-use planners mitigate against further impacts by establishing urban growth limits or boundaries. Effective land management planning relies on dependable, consistent data representing changes in the urban footprint over time.

While many tabular data and spatially-explicit maps currently provide national-scale information on developed lands in the United States (US Department of Agriculture, 2011; Falcone, 2015; Loveland et al., 2002; Mitchell, 1977; Nusser and Goebel, 1997; Price et al., 2006; Sleeter et al., 2013; Soulard et al., 2014; Theobald, 2014; US Census Bureau, 2016a, US Census Bureau, 2016b, US Geological Survey, 2016), few focus on the urban component of developed lands. Developed lands often include both rural and urban areas, and are commonly defined using a modified Anderson classification scheme (Anderson, 1976). Anderson defines developed lands as:

> *Areas of intensive use with much of the land covered by structures. Included in this category are cities, towns, villages, strip developments along highways, transportation, power, and communications facilities, and areas such as those occupied by mills, shopping centers, industrial and commercial complexes, and institutions that may, in some instances, be isolated from urban areas.*

Perhaps the most commonly used resource for maps of developed lands over time is the National Land Cover Database (NLCD). NLCD provides wall-to-wall maps of land cover and land-cover change for the conterminous United States (CONUS) across four map dates (1992, 2001, 2006, and 2011) (Fry et al., 2009; Homer et al., 2004; Jin et al., 2013; Vogelmann et al., 2001; Xian et al., 2009). The 2001, 2006, and 2011 NLCD classification employed classification and regression trees (CART), impervious surface area estimates, and ancillary data to classify land-use/land-cover (LULC) types from Landsat imagery based on a modified Anderson Level II (Anderson 1976) classification, separating developed land into four classes (21-open space, 22-

low-intensity, 23- medium-intensity, and 24-high-intensity) (table 1). According to Homer et al. (2007), these four developed classes are derived by establishing thresholds within NLCD's imperviousness data product. The NLCD developed classes include roads where impervious pixels intersect ancillary vector data from the U.S. Department of Transportation's Bureau of Transportation Statistics (BTS) roads (US Department of Transportation Bureau of Transportation, 2016). The NLCD Level I developed class aggregates the Level II classes 21, 22, 23, and 24 (table 1).

Table 1. USGS National Land Cover Database Level II Developed classes.

NLCD Class	NLCD Class Definition
21	Developed, Open Space - areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
22	Developed, Low Intensity - areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20% to 49% percent of total cover. These areas most commonly include single-family housing units.
23	Developed, Medium Intensity - areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50% to 79% of the total cover. These areas most commonly include single-family housing units.
24	Developed High Intensity - highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80% to 100% of the total cover.

Formal accuracy estimates and a qualitative assessment of NLCD against satellite imagery and aerial photography suggest that NLCD does an effective job mapping developed lands, with Level I user's accuracy ranging from 73-74 % and producer's accuracy ranging from 72-74 % (2001 and 2006, respectively) (Wickham et al., 2013). A characteristic of the NLCD dataset is that rural roads are not differentiated from developed areas that are more traditionally considered part of the urban footprint. This addition of rural roads to core developed areas results in estimates of developed land far higher than the actual area of urban land. Additionally, commission errors of the developed class occur because the road network includes primary and secondary roads that are delineated three or more 30-meter pixels wide, and tertiary roads including gravel tracks and logging routes. Most roads are not wider than 90 meters, which means that NLCD often over-represents actual road area. Road density is inconsistent, with large differences between the western and eastern United States stemming from the use of automated processes in the west to reduce the mapping of dirt tracks through rural areas. Finally, NLCD treats the road network as a static input across the 2001, 2006, and 2011 map dates, unlike the rest of the developed footprint that changes dynamically over time.

While including rural roads as part of the developed footprint may be desirable in some applications focused on examining all impervious lands (Bierwagen et al., 2010; Booth et al., 2002; Powell et al., 2008; Weng, 2012; Yang et al., 2003), end users often prefer a finer level of detail and autonomy to distinguish between urban development, rural development and roads (Brown et al., 2005; Endreny and Thomas, 2009; Hilferink and Rietveld, 1999; Leyk et al., 2014; Theobald, 2014). In select cases, the aforementioned problems with how roads are represented in NLCD have required researchers working on precise land use assessments and reliable empirical models to edit out roads to create quality developed maps as an input (Schwarz et al. 20001; Sleeter et al., 2017; US Environmental Protection Agency, 2017). In many of these assessments and modeling efforts, the proportion of the developed footprint characterized by higher-intensity urban use is of particular interest. To more accurately represent this urban component of developed lands and better understand how urban land changes over time, errors attributable to including rural roads in the NLCD developed class must be resolved.

Other notable, large-scale attempts have been made to improve NLCD developed classes by adding more developed classes and resolving classification problems (Claggett et al., 2013; Falcone, 2015; Theobald, 2014). For example, USGS NAWQA Wall-to-Wall Anthropogenic Land Use Trends (NWALT) is derived from a multi-tiered process incorporating ancillary data to combine land use products and to modify NLCD by thinning rural roads (Falcone, 2015). These modified products fix some errors associated with the inclusion of roads in the NLCD products (Endreny and Thomas, 2009) but neither provide a comprehensive, consistent product representative of urban land cover with corresponding accuracy estimates, nor describe a benchmark method that can be replicated.

The objective of our study is to generate more reliable maps of urban lands in the United States by removing rural roads from the NLCD developed class for each map date while retaining developed lands represented by urban infrastructure and urban roads. The road removal process relies on merging NLCD developed maps with readily available geospatial data, applying multiple filtering processes, and performing automated and manual editing to correct the urban footprint. We create urban maps for CONUS for four dates (1992, 2001, 2006, and 2011). Finally, an accuracy assessment is performed on the newly created urban maps and the NLCD maps for the years 2001 and 2006 to evaluate how well the different products capture urban lands. The new urban maps provide accurate spatially explicit data on urban land and urban change critical to research on the causes and consequences of urban expansion across the country.

2. Materials and Methods

2.1 Removing rural roads from NLCD

In general, the removal of rural roads from the NLCD developed classes was achieved by applying several neighborhood filtering operations and various manual and automated editing procedures to resolve omission and commission errors (fig. 1). We did not attempt to remove roads from urban areas because roads within high-density development are considered part of the urban class. The first step in the process was to aggregate the 2001 NLCD (Homer et al., 2007) Level II developed classes into just one class (Level I developed) to provide a base from which to identify urban areas (Level I and Level II are based on Anderson (1976)). Unlike the 1992 NLCD map (Vogelmann et al., 2001), the 2001 NLCD map is the first year that roads were included in the developed classes. The Level I developed class provided an ideal starting point for delineating urban areas because the four Level II developed classes coalesce into larger clumps or contiguous groups of pixels when viewed as one thematic layer (fig. 2A). Many of these clumps were clearly urban areas where roads, residential, commercial, and industrial areas had merged. The NLCD Level I developed class was then intersected with a circa-2000 BTS roads dataset (US Department of Transportation Bureau of Transportation, 2016) to identify which NLCD developed pixels corresponded to roads defined by BTS, and to include roads omitted by NLCD in the western United States. The fusion of NLCD and BTS data also ensured that subsequent process steps were applied uniformly across the country, and that implementation of our methods did not result in the loss of small urban developments throughout the West.

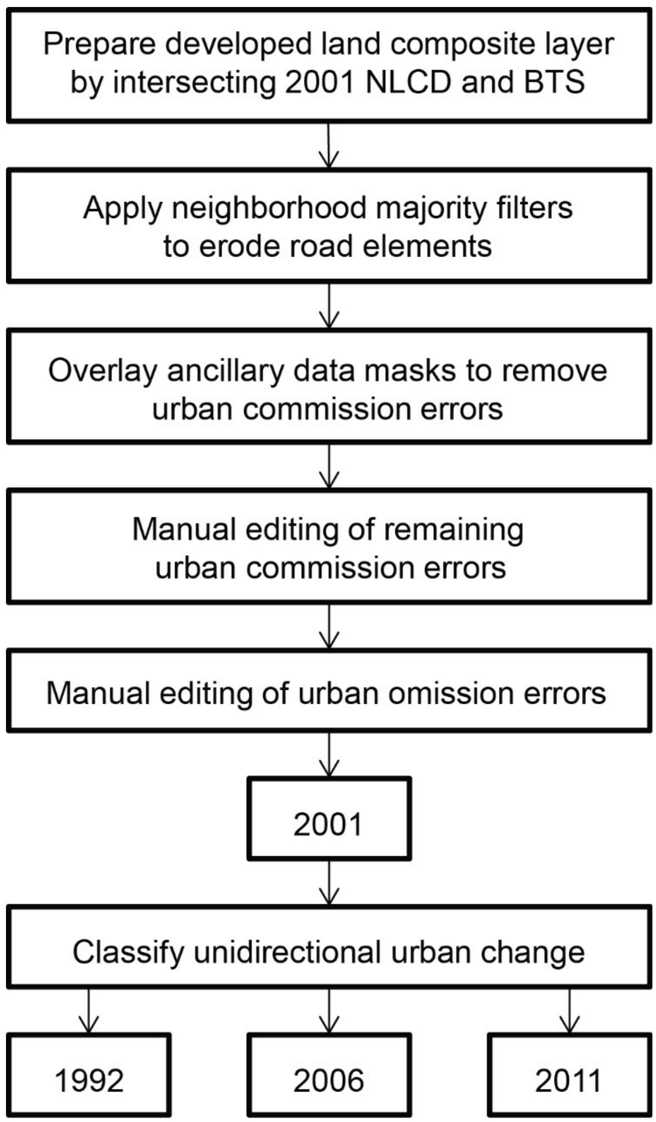

Figure 1 (black/white): Conceptual diagram describing the process of removing roads from NLCD to create urban maps. The process includes neighborhood filtering, as well as automated and manual editing processes to resolve commission and omission errors.

The aggregated NLCD 2001 and BTS composite was then used as the starting point for identifying urban areas by removing pixels that corresponded to rural roads. We evaluated a variety of filtering functions to remove road pixels. We initially tested small neighborhood filters to erode the road pixels in the NLCD/BTS composite (fig. 3). Neighborhood filters are morphological operators available in the ERDAS Imagine (Intergraph, 2014) software package that use a square structuring element (also known as a kernel or array) that passes over each cell in the classified map and performs an operation based on the characteristics of the neighboring pixels. For example, a 3x3 majority filter recodes the center pixel in the kernel to the value corresponding to the class that represents the majority of pixels in the larger 9-cell array. While small kernels (3x3) effectively removed scattered pixels and narrow road elements, they were ineffective at removing wide road elements. After testing four different kernel sizes, we settled on a 9x9 majority filter because it removed most road pixels without excessively eroding core urban areas. Next a second filter pass was applied to remove developed pixels remaining where roads in the NLCD/BTS composite were widest. Collectively, this process removed small isolated pixels and thin linear features in the developed image and left large contiguous groups of pixels intact. To restore urban pixels omitted after the initial filtering was complete, we performed an inverse filtering function (also referred to as dilation) to re-insert developed pixels along the edge of urban areas that had been nibbled away during removal of road pixels. A qualitative assessment of the filtered NLCD/BTS composite suggested that the preceding filtering steps provided an improved delineation of urban areas, but problems remained. The filtered product incorrectly retained groups of pixels (i.e., commission errors) that corresponded to wide sections of primary roads, highway interchanges, road intersections with adjacent development, and sections of highways with parallel frontage roads. Moreover, the filtered product incorrectly removed groups of pixels (i.e., omission errors) where urban patches were too small to be retained as defined by the filtering algorithm.

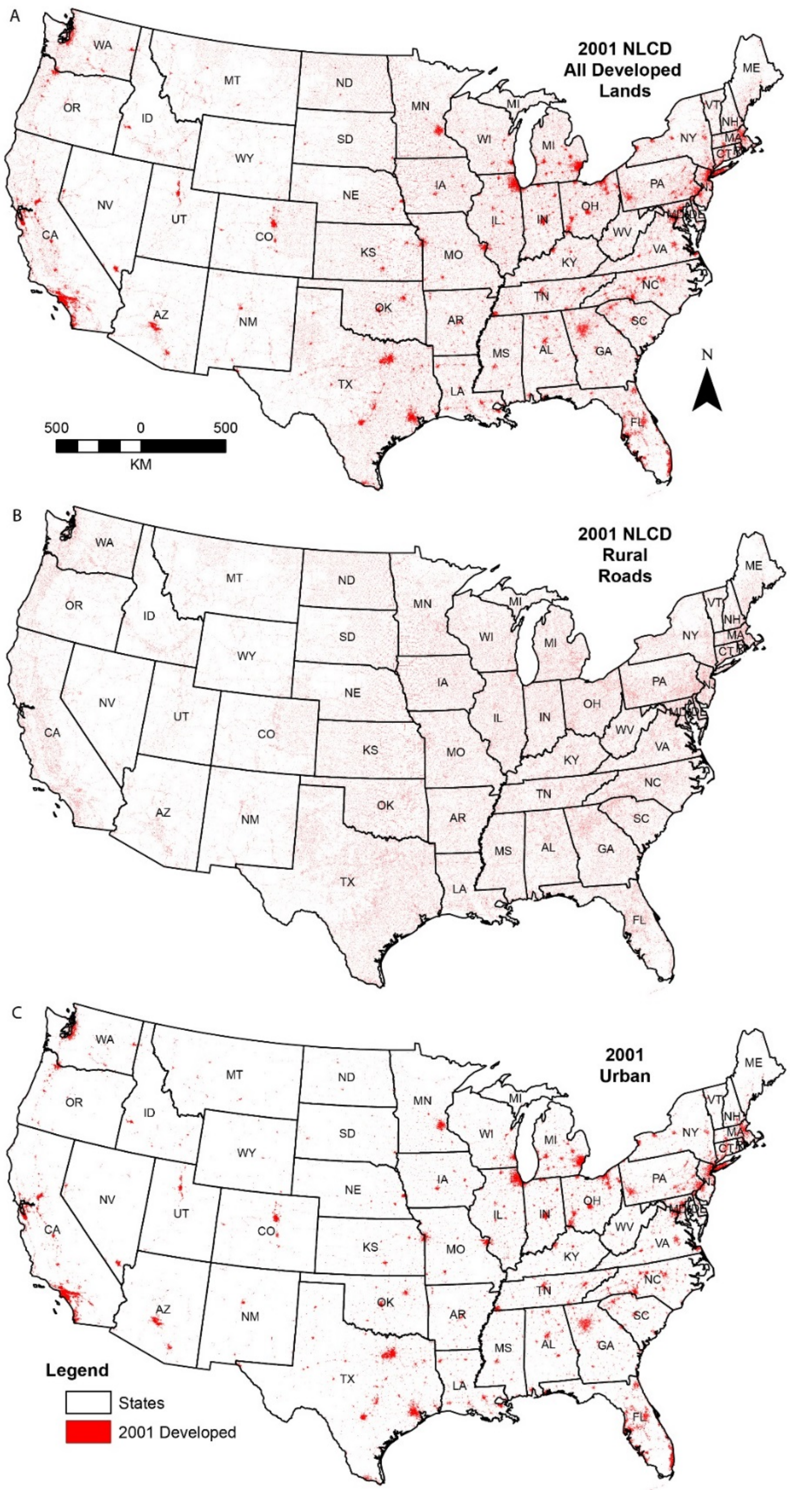

Figure 2 (color): Illustration representing (A) lands mapped as developed by 2001 NLCD, (B) rural roads removed from NLCD developed lands, and (C) 2001 urban map where speckle associated with roads has been removed.

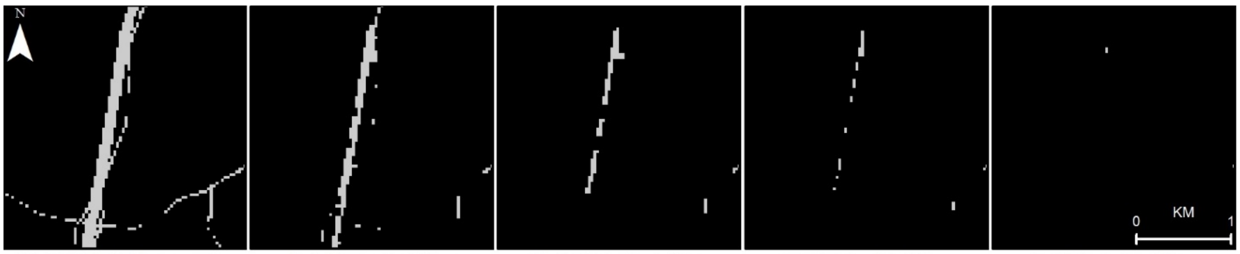

Figure 3 (grayscale). Examples of different kernel filters applied to the NLCD/BTS composite near Butte, MT (45°52'55.2"N, 112°40'22.8"W) indicates that smaller filters were insufficient in removing road pixels (gray).

To improve the filtered product and to remove pixels not considered part of the urban footprint, we performed a clump function using ERDAS IMAGINE software (Intergraph, 2014) and deleted clumps using various criteria. The filtered product was converted to polygon features and small clumps with fewer than 100 pixels that did not intersect with point features in the Geographic Names Information System (GNIS) populated places (US Geological Survey, 2013) were deleted. Larger clumps (equal to or greater than 1,250 pixels) and clumps that corresponded to the GNIS database were retained as urban areas. All clumps larger than 100 pixels but less than 1,250 pixels in size were screened using a manual editing process. Long linear clumps designated as urban areas were visually inspected using Google Earth's historical imagery and manually deleted if determined to be a rural road.

Developed pixels incorrectly removed during filtering were re-inserted into the urban footprint using a combination of masks and manual editing steps. Pixels removed during filtering that intersected with the BTS (US Department of Transportation Bureau of Transportation, 2016) or National Overview Road Metrics (NORM) Euclidean distance (ED) (Watts, 2005) datasets were considered correctly classified non-urban areas. However, pixels that did not intersect these datasets presented a challenging editing problem because they included valid urban areas that needed to be restored, but also included areas not considered part of the urban class, including misaligned roads, road stubs, road shoulders, railroads, portions of fallow fields, drilling pads, and patches of barren. Masks were compiled by using proximity buffers around roads, railroads, and mining locations (Soulard et al., 2016; US Geological Survey, 2016). Pixels that either fell within 90 meters of roads and railroads or that intersected mine boundaries were classified as "not urban". Finally, a human interpreter performed a manual editing step to correct any remaining easily visible errors (fig. 4).

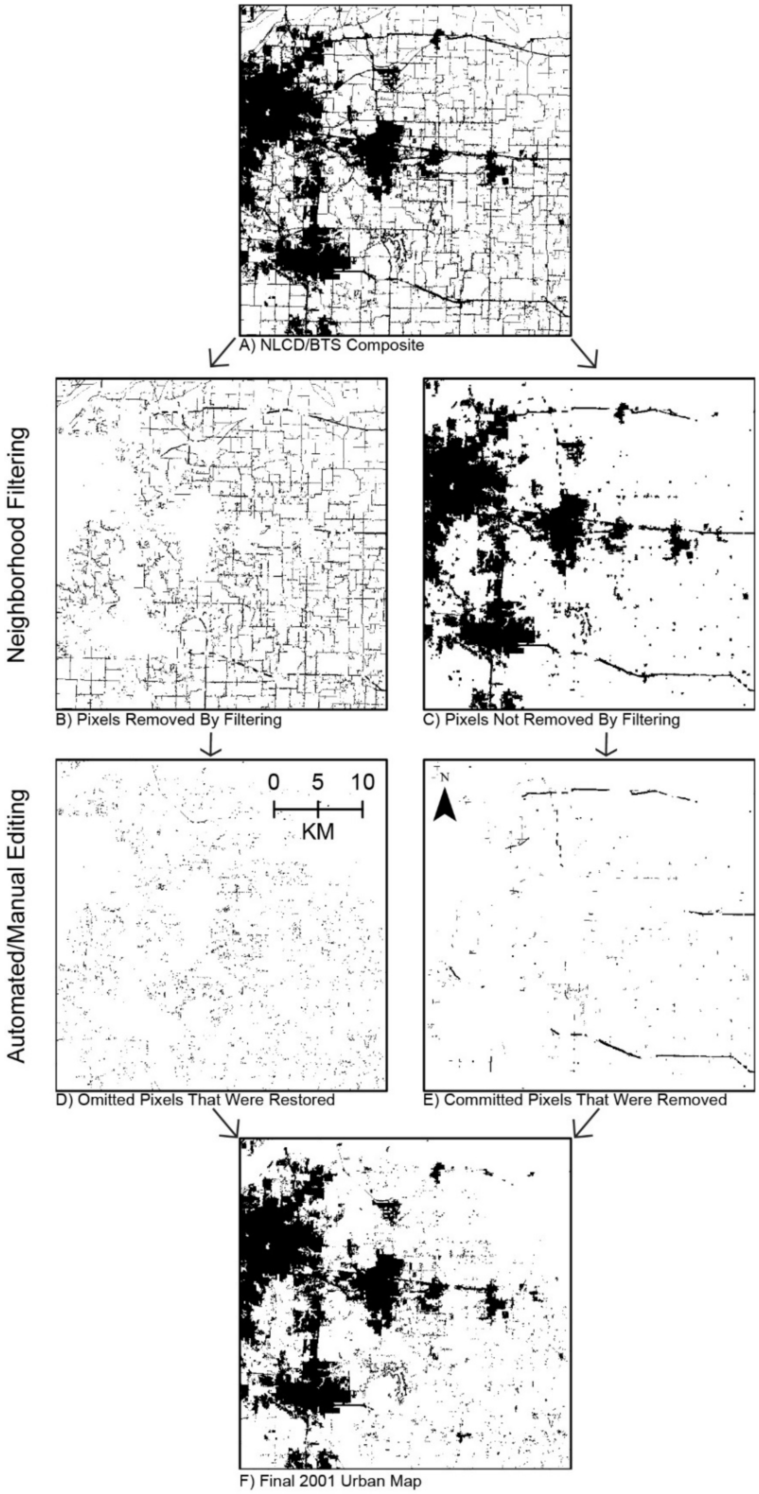

Figure 4 (grayscale). Map of Kansas City, MO illustrating how automated and manual editing processes were applied to resolve commission and omission errors remaining after the filtering steps.

2.2 Mapping urban change

Once the road removal and clean-up of the 2001 urban footprint were complete (fig. 2B and 2C), we evaluated the developed pixels added between NLCD 2001 and NLCD 2006. Only conversions "to developed" were considered in this step, since we assume changes in developed lands are unidirectional. New roads that were built between 2001 and 2006 were manually deleted, along with oil and gas drilling pads classified as developed. The process was repeated for NLCD developed pixels added between 2006 and 2011. New roads that were built between 2006 and 2011 were deleted, as well as oil and gas drilling pads classified as developed. Different procedures were used to create the 1992 urban footprint. A composite of the disaggregated 2001 developed layer, NLCD 1992 developed (Vogelmann et al., 2001), NLCD Retrofit 1992 developed (Fry et al., 2009), and NWALT 1992 (Falcone, 2015) was created by intersecting all layers and comparing pixels mapped in all 1992 products with the developed pixels in the 2001 product. Urban areas were back-classified to 1992 by retaining only those pixels from NLCD 1992, NLCD Retrofit 1992, or NWALT 1992 that were located within the 2001 urban footprint. This process ensured that development followed a unidirectional conversion process.

The urban (see Supplementary Material) and urban change maps were summarized nationally for illustrative purposes. Regional results were calculated by intersecting the original NLCD maps and the final urban maps with Level III EPA ecoregions (Omernik, 1987). Comparative maps and statistics were compiled to highlight key differences between each set of products.

2.3 2001 and 2006 Accuracy Assessment

The accuracy of the final versions of the 2001 and 2006 maps of the binary classification "urban" and "not urban" was evaluated where rural roads were included in the "not urban" class. The sample selected for the accuracy assessment of the 2006 NLCD products (Wickham et al., 2013) provided the sample locations for assessing the accuracy of the urban maps. The NLCD sampling design was stratified geographically by ten regions of the United States, and within each geographic stratum further stratified by land-cover change. For each sampled pixel, reference labels were obtained for 2001 and 2006 (the interpreter did not know the map class label). The original NLCD reference classification provided from Wickham et al. (2013) for each of the 15,000 sample pixels included a primary and alternate classification for each sample pixel. Because "urban" is not one of the NLCD land-cover classes, the NLCD sample pixels had to be re-interpreted to assess accuracy of the urban class. Given that the urban class is a subset of the NLCD developed class, the re-interpretation of the reference class labeling could be limited to only those sample pixels that had a primary or alternate label of "developed" in the NLCD reference sample database.

For the 2001 reference classification we used circa-2001 Landsat imagery and 1998–2001 aerial imagery to determine if the pixel should be re-labeled from "developed" to "not urban" if the pixel was a rural road or fell more than 500 meters outside of a designated urban area (US Census Bureau, 2016b). The 2006 reference sample pixels were similarly re-examined using 2005 Landsat imagery and 2003–2007 aerial imagery to re-label sample pixels from "developed" to "not urban" if the pixel was a rural road or was located well outside of an urban area. Through this procedure we effectively revised the definition of the remaining developed reference sample pixels to correspond to our definition of "urban". Our reference data interpretation protocol mirrors that of the NLCD accuracy assessment in that a 3-by-3 pixel window was applied to provide context for the class labeling (Wickham et al., 2010). Accuracy estimates and standard errors were produced using the same formulas reported by Wickham et al. (2013).

3. Results

3.1 Comparative Assessment of NLCD and Urban Maps

In the NLCD products for 2001, 2006, and 2011, the developed class totaled 434,990 km^2, 446,660 km^2, and 454,290 km^2 respectively. The mapped developed area represents roughly 6% of the land area in the conterminous United States. The change in map area of the NLCD developed class was 12,670 km^2 for 2001-2006 and 7,630 km^2 for 2006-2011. The process of creating urban maps by removing rural roads from the NLCD developed class resulted in the deletion of 230,360 km^2, 230,670 km^2, and 230,880 km^2 from the NLCD developed class area in 2001, 2006, and 2011, respectively. The resulting urban maps included far less area of urban land, totaling 173,580 km^2 in 1992, 203,640 km^2 in 2001, 215,990 km^2 in 2006, and 223,410 km^2 in 2011, or about 3% of the land area in the conterminous United States (see Supplementary Material). The annual rate of change in urban area decreased over time. Specifically, the change in mapped area of urban land was 3,339 km^2 per year for 1992-2001 (30,050 km^2 total), 2,472 km^2 per year for 2001-2006 (12,360 km^2 total), and 1,484 km^2 per year for 2006-2011 (7,420 km^2 total).

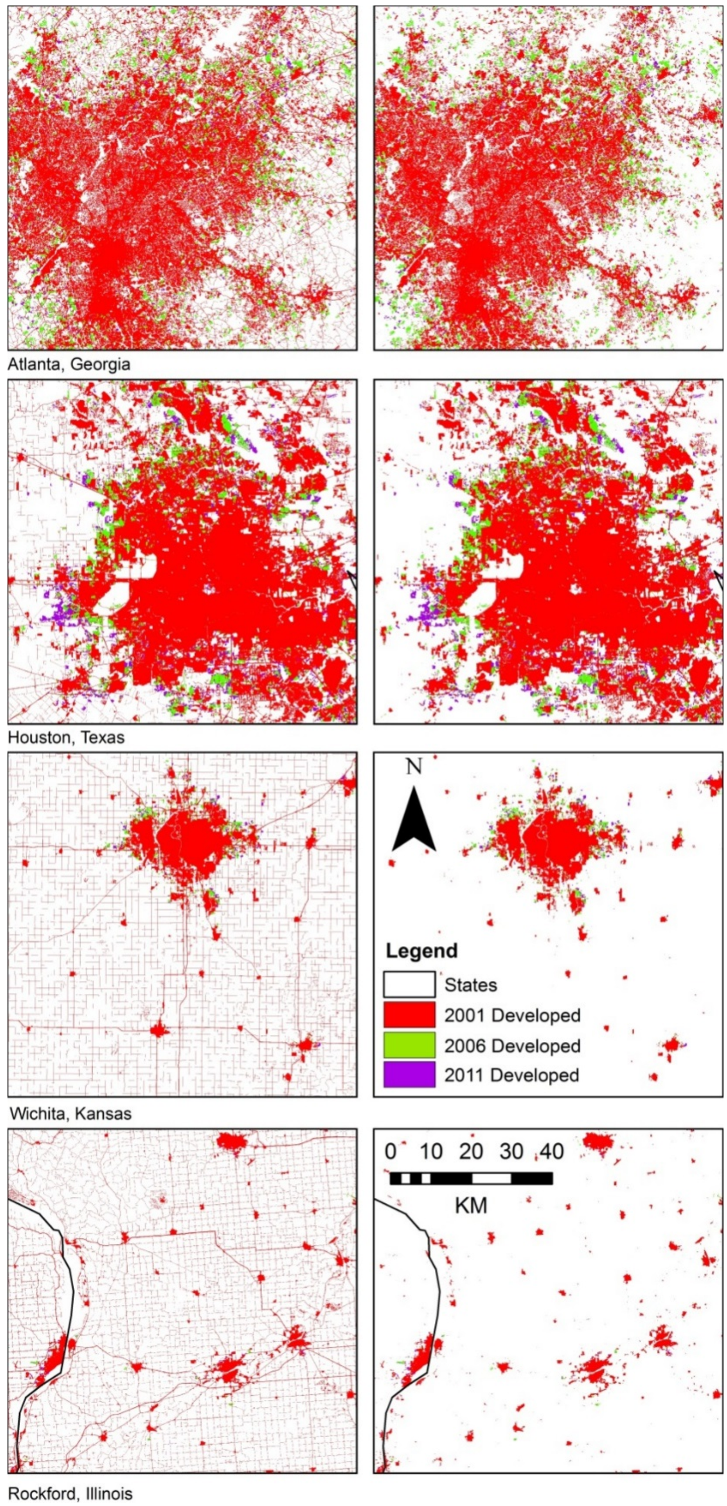

Figure 5 (color): Comparisons between NLCD developed class (left panel) and the new urban maps (right panel) for four urban areas in the United States. Three dates are shown (2001, 2006, and 2011).

Within Level III Ecoregions the total area of NLCD developed land differs substantially from the total area of urban land (fig. 5; table 2). For example, many of the ecoregions with the largest developed footprint in NLCD 2011 also contain a sizable urban footprint in the 2011 urban map. The top 10 ecoregions of total NLCD developed area and the top 10 of total urban area include ecoregions in the eastern and western United States with major urban centers (table 2): Ecoregion 75 includes major urban centers throughout northern and central Florida including Jacksonville, Orlando, and St. Petersburg, Ecoregion 45 includes Atlanta (Georgia), Charlotte (North Carolina) and Raleigh-Durham (North Carolina), Ecoregion 6 includes the San Francisco-Bay Area and Los Angeles-San Diego Area (California), and Ecoregion 67, which is the second largest ecoregion in the United States, includes Mobile (Alabama), Tallahassee (Florida), Fayetteville (North Carolina), Richmond (Virginia), and substantial portions of Baltimore (Maryland) and Washington (District of Columbia). Although total areas and rankings differ between the two maps, seven of the ecoregions with the largest developed footprint in NLCD 2011 are also included in the urban map list. The most notable differences between NLCD and the urban maps are that Ecoregions 47 (Western Corn Belt), 27 (Central Great Plains), and 25 (Western High Plains) are no longer among the top ten. The smaller urban land area relative to the developed land area across these ecoregions is the result of removal of rural roads during editing. In the ecoregions removed from the top 10 of total developed area, the only major urban centers are Omaha (Nebraska) and Des Moines (Iowa) in Ecoregion 47, Wichita (Kansas) and Oklahoma City (Oklahoma) in Ecoregion 27, and Denver (Colorado) in Ecoregion 25.

Table 2. Level III Ecoregions with the greatest total area of developed land in 2011 NLCD and the greatest total area of urban land in the 2011 urban map.

Ecoregion Name	Eco #	NLCD 2011 Developed Area (km²)
Southeast Plains	65	24,740
Piedmont	45	21,706
South Coastal Plain	75	18,144
West Corn Belt Plains	47	15,918
CA Oak Woodlands	6	15,817
Central Great Plains	27	14,746
Ridge and Valley	67	13,883
Central Corn Belt Plains	54	13,656
West High Plains	25	12,927
East Corn Belt Plains	55	11,625

Ecoregion Name	Eco #	Urban 2011 Area (km²)
South Coastal Plain	75	14,041
Piedmont	45	13,700
CA Oak Woodlands	6	12,630
Southeast Plains	65	10,194
Central Corn Belt Plains	54	8,799
Northeast Coastal Zone	59	8,470
Ridge and Valley	67	7,795
MI/IN Drift Plains	56	7,593
East Corn Belt Plains	55	7,310
North Piedmont	64	6,678

3.2 2001 and 2006 Accuracy Assessment Results

Both the urban and NLCD maps for 2001 and 2006 exceeded the overall accuracy benchmark of 85% (Anderson, 1976) for the binary classification of urban and not urban. The overall accuracy of the urban classification improved following removal of rural roads from the NLCD developed class, primarily due to the substantial reduction in urban commission errors (table 3). Over 230,000 km^2 of area formerly classified as developed by the NLCD was removed to create the new urban maps resulting in an increase in overall accuracy of 2.5% and a reduction in urban commission error by over 30% relative to the NLCD developed class. The improvement in urban commission error was achieved at the small expense of increasing the urban omission error by 4% relative to the NLCD developed class (table 3).

Table 3. Confusion matrix for 2001 and 2006 NLCD and urban maps. Values shown are percent area of the conterminous United States. User's and producer's accuracies, as well as commission and omission error estimates, are reported as percents with standard errors in parentheses.

2001 NLCD Developed Map

NLCD Developed Map

	Reference				
	Not Urban	Urban	Total	User	Commission
Not Urban	93.93	0.48	94.42	99 (1)	1 (1)
Urban	3.32	2.27	5.58	41 (2)	59 (2)
Total	97.25	2.75			
Producer	97 (1)	82 (5)			
Omission	3 (1)	18 (5)		Overall	96.2% (0.2)

2001 Urban Map

Urban Map

	Reference				
	Not Urban	Urban	Total	User	Commission
Not Urban	96.56	0.60	97.16	99 (1)	1 (1)
Urban	0.69	2.15	2.84	76 (2)	24 (2)
Total	97.25	2.75			
Producer Accuracy	99 (1)	78 (5)			
Omission	1 (1)	22 (5)		Overall	98.7% (0.2)

2006 NLCD Developed Map

NLCD Developed Map

	Reference				
	Not Urban	Urban	Total	User	Commission
Not Urban	93.80	0.46	94.26	99 (1)	1 (1)
Urban	3.30	2.44	5.74	43 (2)	57 (2)
Total	97.10	2.90			

Producer	97 (1)	84 (5)				
Omission	3 (1)	16 (5)			Overall	96.2% (0.2)

2006 Urban Map

Reference

Urban Map		Not Urban	Urban	Total	User	Commission
	No Urban	96.42	0.59	97.01	99 (1)	1 (1)
	Urban	0.68	2.31	2.99	77 (2)	23 (2)
	Total	97.10	2.90			
	Producer	99 (1)	80 (5)			
	Omission	1 (1)	20 (5)		Overall	98.7% (0.2)

Examining the NLCD and urban maps at the 692 re-interpreted ground reference sample points found along rural roads indicates that a higher percent of developed land mapped by NLCD in 2001 and 2006 intersects rural roads (40% and 46%, respectively) than the 2001 and 2006 urban maps (3% and 7%, respectively). Only 23 rural road reference sample pixels were misclassified in the 2001 urban maps, and only 50 rural road reference sample pixels remained misclassified in 2006.

Discussion

Trends in area of developed land reported by different studies show considerable variability (fig. 6). The inclusion of roads in the developed class of NLCD and the way that BTS road data were used to map developed land result in much higher areas of developed land in NLCD relative to USGS Land Cover Trends (LCT) (Sleeter et al., 2013) data, the USDA National Resource Inventory (NRI) (US Department of Agriculture, 2013), the USDA Economic

Research Service (ERS) data (US Department of Agriculture, 2011), and NWALT dataset (Falcone, 2015) (fig. 6). The accuracy assessment results using the urban reference class information demonstrate that NLCD includes substantial non-urban area in its developed class across CONUS, as evidenced by high urban commission errors in 2001 and 2006 (table 3). By removing approximately 230,000 km^2 of rural road area from the NLCD developed class for each of the three dates mapped, we created urban maps with substantially smaller area of urban relative to the total area of the NLCD developed class. The urban maps also had smaller area than other estimates provided by map and survey data generated by USGS and USDA (fig. 6). By excluding rural roads and small patches of rural development from NLCD, we created urban maps that more accurately portray the true core urban areas and consequently serve to better monitor the rates of change and spatial patterns of dense urban clusters.

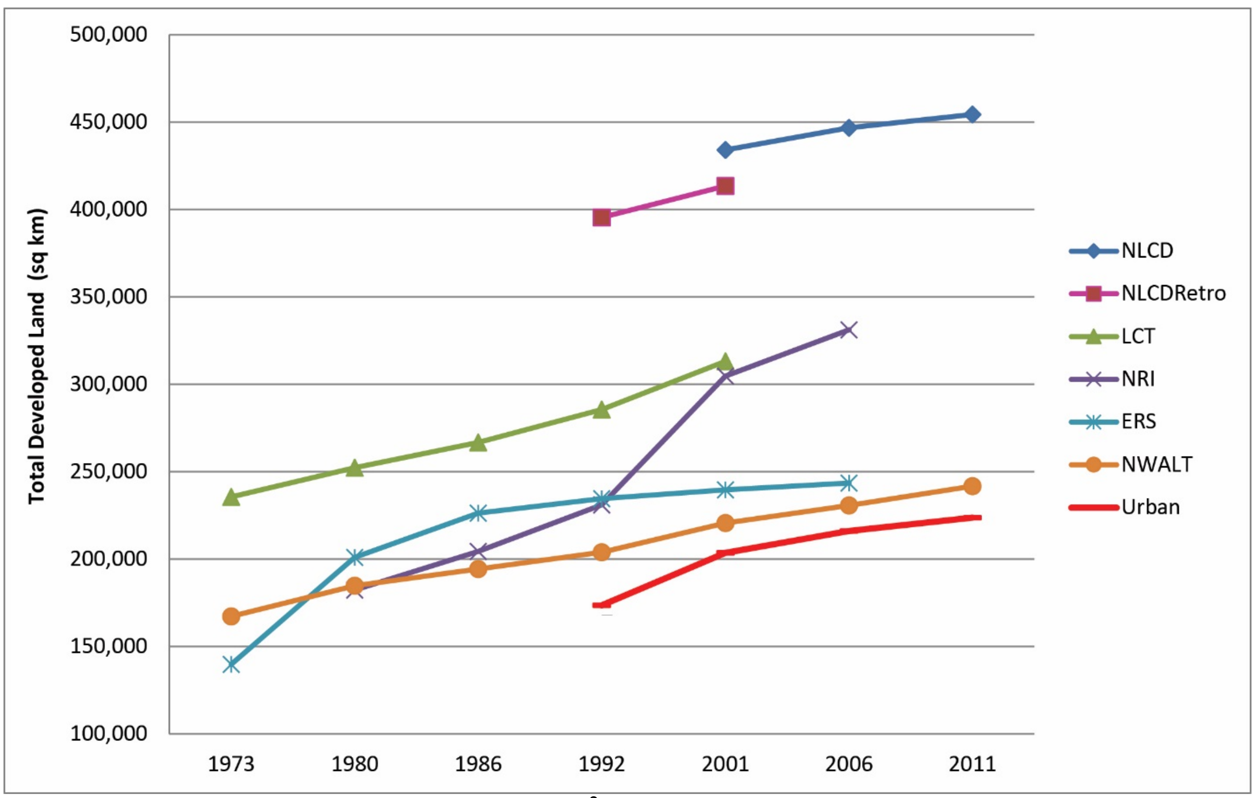

Figure 6 (color): Area of developed land (km^2) for CONUS across all readily accessible tabular

data and spatially-explicit maps. The Urban results represent the maps created by removing rural roads from the NLCD developed class.

The urban maps reduced urban commission errors by over 30% relative to NLCD while trading off only a 4% increase in urban omission error. Further, revising the NLCD maps led to substantially fewer roads in the urban maps. Comparing NLCD and urban maps based on the reference sample data illustrates that the processing removed 92% of rural roads in 2001 and 84% of rural roads in 2006. Lower commission errors relative to NLCD provide evidence that NLCD includes roads that should not be classified as urban land use. On the other hand, higher urban omission errors are the result of the multi-tiered clean-up processes, where the smaller urban footprint of the new maps omits some real urban areas in the process of removing roads. The removal of rural roads in the NLCD developed layers results in an opportunity to incorporate contemporary road data compiled using advanced methods, such as TIGER (U.S. Census Bureau, 2017) or private sector data from HERE (formerly NAVTEQ) (HERE, 2016), into the developed class.

Conclusions

Early in our exploratory effort to review readily available information on urban lands in the United States, we determined that rural roads included in the developed class of NLCD maps were problematic due to inconsistencies in road location, density, and continuity. NLCD also treats the road network as a static input. The process described in this article provides a way to modify the NLCD to map an “urban” class by reducing errors associated with mapping rural roads as “developed”. The benefits afforded by removing errors associated with the mapping of roads and more effectively mapping urban lands from 1992 to 2011 are far reaching. Our mapping of urban lands not only leads to more accurate results but also allows for characterizing actual urban growth rather than assessing changes for a more generalized developed class that includes rural roads and other small patches of rural developed land. Based on our results, the following conclusions can be made.

- The removal of approximately 230,000 km^2 of rural roads from the NLCD developed classes led to substantially less urban land area relative to the NLCD developed land area for each map date and this difference in area likely will impact modeling results and policy decisions for users of these data.
- The urban maps have higher overall accuracy, much lower urban commission error rates, and only slightly higher urban omission error rates relative to the NLCD developed class.
- Most pixels removed from the NLCD developed class and labeled as Not Urban were located in the eastern US within regions characterized by low-density development and an abundance of rural roads.

Reliable maps of urban change are essential to further understanding of the causes and consequences of human land use practices on the landscape. The urban maps produced by removing rural roads as well as other small areas of rural development from the NLCD developed class refine the characterization of urban development patterns and change rates and this information will assist scientists, planners, and land managers to better understand and quantify driving forces and impacts of urban change.

Acknowledgments

To be completed after blind review.

Author Contributions

To be completed after blind review.

Supplementary Material

Binary digital maps of urban land in the United States for 1992, 2001, 2006, and 2011 will be provided through the journal as reduced spatial resolution (90-meter) Erdas Imagine (IMG) files. The projection is set to Albers Equal-Area Conic, North American Datum of 1983. Full-resolution (30-meter) digital maps will be available upon publication [to be completed after blind review].

References

Anderson, J.R., 1976. A land use and land cover classification system for use with remote sensor data. US Government Printing Office.

Bengston, D.N., Fletcher, J.O., Nelson, K.C., 2004. Public policies for managing urban growth and protecting open space: policy instruments and lessons learned in the United States. Landscape and Urban Planning 69, 271–286. doi:10.1016/j.landurbplan.2003.08.007

Bierwagen, B.G., Theobald, D.M., Pyke, C.R., Choate, A., Groth, P., Thomas, J.V., Morefield, P., 2010. National housing and impervious surface scenarios for integrated climate impact assessments. Proceedings of the National Academy of Sciences 107, 20887–20892.

Booth, D.B., Hartley, D., Jackson, R., 2002. Forest cover, impervious-surface area, and the mitigation of stormwater impacts. JAWRA Journal of the American Water Resources Association 38, 835–845.

Brown, D.G., Johnson, K.M., Loveland, T.R., Theobald, D.M., 2005. Rural land-use trends in the conterminous United States, 1950-2000. Ecological Applications 15, 1851–1863.

Claggett, P.R., Irani, F.M., Thompson, R.L., 2013. Estimating the extent of impervious surfaces and turf grass across large regions. JAWRA Journal of the American Water Resources Association 49, 1057-1077. doi:10.1111/jawr.12110

Endreny, T.A., Thomas, K.E., 2009. Improving estimates of simulated runoff quality and quantity using road-enhanced land cover data. Journal of Hydrologic Engineering 14, 346–351.

Fahrig, L., 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics 34, 487–515.

Falcone, J.A., 2015. US Conterminous Wall-to-Wall Anthropogenic Land Use Trends (NWALT), 1974–2012. U.S. Geological Survey.

Fry, J.A., Coan, M.J., Homer, C.G., Meyer, D.K., Wickham, J.D., 2009. Completion of the National Land Cover Database (NLCD) 1992-2001 Land Cover Change Retrofit Product (No. 2008–1379). U.S. Geological Survey.

Göbel, P., Stubbe, H., Weinert, M., Zimmermann, J., Fach, S., Dierkes, C., Kories, H., Messer, J., Mertsch, V., Geiger, W.F., Coldewey, W.G., 2004. Near-natural stormwater management and its effects on the water budget and groundwater surface in urban areas taking account of the hydrogeological conditions. Journal of Hydrology 299, 267–283. doi:10.1016/j.jhydrol.2004.08.013

Hilferink, M., Rietveld, P., 1999. Land Use Scanner: An integrated GIS based model for long term projections of land use in urban and rural areas. Journal of Geographical Systems 1, 155–177.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., Wickham, J., others, 2007. Completion of the 2001 national land cover database for the conterminous United States. Photogrammetric Engineering and Remote Sensing 73, 337.

Homer, C., Huang, C., Yang, L., Wylie, B., Coan, M., 2004. Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering & Remote Sensing 70, 829–840.

Intergraph, 2014. Erdas Imagine 2014. Hexagon Geospatial.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., Xian, G., 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment 132, 159–175. doi:10.1016/j.rse.2013.01.012

Leyk, S., Ruther, M., Buttenfield, B.P., Nagle, N.N., Stum, A.K., 2014. Modeling residential developed land in rural areas: A size-restricted approach using parcel data. Applied Geography 47, 33–45.

Loveland, T.R., Sohl, T.L., Stehman, S.V., Gallant, A.L., Sayler, K.L., Napton, D.E., 2002. A strategy for estimating the rates of recent United States land cover changes. Photogrammetric Engineering & Remote Sensing 68, 1091–1099.

Mitchell, W.B., 1977. GIRAS: A geographic information retrieval and analysis system for handling land use and land cover data. U.S. Government Printing Office.

Nusser, S.M., Goebel, J.J., 1997. The National Resources Inventory: a long-term multi-resource monitoring programme. Environmental and Ecological Statistics 4, 181–204.

Omernik, J.M., 1987. Ecoregions of the Conterminous United States. Annals of the Association of American Geographers 77, 118–125. doi:10.1111/j.1467-8306.1987.tb00149.x

Powell, S.L., Cohen, W.B., Yang, Z., Pierce, J.D., Alberti, M., 2008. Quantification of impervious surface in the Snohomish water resources inventory area of western Washington from 1972–2006. Remote Sensing of Environment 112, 1895–1908.

Schwarz, G.E., Smith, R.A., Alexander, R.B. and Gray, J.R., 2001. A spatially referenced regression model (SPARROW) for suspended sediment in streams of the conterminous US. In 7th Federal Interagency Sedimentation Conference.

Sleeter, B.M., Sohl, T.L., Loveland, T.R., Auch, R.F., Acevedo, W., Drummond, M.A., Sayler, K.L., Stehman, S.V., 2013. Land-cover change in the conterminous United States from 1973 to 2000. Global Environmental Change 23, 733–748. doi:10.1016/j.gloenvcha.2013.03.006

Sleeter, B.M., Wood, N.J., Soulard, C.E. and Wilson, T.S., 2017. Projecting community changes in hazard exposure to support long-term risk reduction: A case study of tsunami hazards in the US Pacific Northwest. International Journal of Disaster Risk Reduction, 22, pp.10-22.

Soulard, C.E., Acevedo, W., Auch, R.F., Sohl, T.L., Drummond, M.A., Sleeter, B.M., Sorenson, D.G., Kambly, S., Wilson, T.S., Taylor, J.L., Sayler, K.L., 2014. Land Cover Trends Dataset, 1973–2000 (Data Series No. 844). USGS.

Soulard, C.E., Acevedo, W., Stehman, S.V., Parker, O.P., 2016. Mapping extent and change in surface mines within the United States for 2001 to 2006. Land Degradation & Development 27, 248–257. doi:10.1002/ldr.2412

Strauch, G., Möder, M., Wennrich, R., Osenbrück, K., Gläser, H.-R., Schladitz, T., Müller, C., Schirmer, K., Reinstorf, F., Schirmer, M., 2008. Indicators for assessing anthropogenic impact on urban surface and groundwater. Journal of Soils and Sediments 8, 23–33. doi:10.1065/jss2007.06.234

Swenson, J.J., Franklin, J., 2000. The effects of future urban development on habitat fragmentation in the Santa Monica Mountains. Landscape Ecology 15, 713–730.

Taha, H., Akbari, H., Rosenfeld, A., Huang, J., 1988. Residential cooling loads and the urban heat island—the effects of albedo. Building and Environment 23, 271–283.

Theobald, D.M., 2014. Development and applications of a comprehensive land use classification and map for the US. PLoS ONE 9, e94628. doi:10.1371/journal.pone.0094628

US Census Bureau, 2016a. American FactFinder [WWW Document]. URL http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml (accessed 6.8.16).

US Census Bureau, 2016b. Cartographic Boundary Shapefiles - Urban Areas [WWW Document]. URL https://www.census.gov/geo/maps-data/data/cbf/cbf_ua.html (accessed 6.8.16).

US Census Bureau, 2017. TIGER/Line® Shapefiles and TIGER/Line® Files [WWW Document]. URL https://www.census.gov/geo/maps-data/data/tiger-line.html (accessed 6.8.16).

US Department of Agriculture, 2013. Summary Report: 2010 National Resources Inventory. Washington, DC, and Center for Survey Statistics and Methodology, Iowa State University, Ames, Iowa: Natural Resources Conservation Service.

US Department of Agriculture, 2011. Major Land Uses. U.S. Department of Agriculture, Economic Research Service, URL https://www.ers.usda.gov/data-products/major-land-uses/major-land-uses/#Urban area (accessed 6.8.16).

US Department of Transportation Bureau of Transportation, 2016. Geospatial Information [WWW Document]. URL http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/subject_areas/geographic_information_services/index.html (accessed 6.8.16).

US Environmental Protection Agency, 2017. EnviroAtlas - Dasymetric Population for the Conterminous United States. URL http://enviroatlas.epa.gov/enviroatlas/ (accessed 10.5.17).

US Geological Survey, 2013. The National Map: Cities and Towns of the United States [WWW Document]. URL http://nationalmap.gov/small_scale/mld/1cities.html (accessed 6.8.16).

US Geological Survey, 2016a. One Million-Scale Railroads of the United States [WWW Document]. URL http://nationalmap.gov/small_scale/mld/1rails.html (accessed 6.8.16).

US Geological Survey, 2016b. National Gap Analysis Program (GAP) Land Cover Data Portal [WWW Document]. URL http://gapanalysis.usgs.gov/gaplandcover/data/download/ (accessed 6.8.16).

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, N., 2001. Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper Data and Ancillary Sources. Photogrammetric Engineering and Remote Sensing 67, 650–662.

Wang, J., Chen, Y., Shao, X., Zhang, Y., Cao, Y., 2012. Land-use changes and policy dimension driving forces in China: Present, trend and future. Land Use Policy 29, 737–749. doi:10.1016/j.landusepol.2011.11.010

Watts, R.D., 2005. Distance to nearest road in the conterminous United States (Fact Sheet), 2005-3011. U.S. Geological Survey.

Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. Remote Sensing of Environment 117, 34–49.

Wickham, J.D., Stehman, S.V., Fry, J.A., Smith, J.H., Homer, C.G., 2010. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. Remote Sensing of Environment 114, 1286–1296. doi:10.1016/j.rse.2010.01.018

Wickham, J.D., Stehman, S.V., Gass, L., Dewitz, J., Fry, J.A., Wade, T.G., 2013. Accuracy assessment of NLCD 2006 land cover and impervious surface. Remote Sensing of Environment 130, 294–304. doi:10.1016/j.rse.2012.12.001

Xian, G., Homer, C., Fry, J., 2009. Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. Remote Sensing of Environment 113, 1133–1147. doi:10.1016/j.rse.2009.02.004

Yang, L., Xian, G., Klaver, J.M., Deal, B., 2003. Urban land-cover change detection through sub-pixel imperviousness mapping using remotely sensed data. Photogrammetric Engineering & Remote Sensing 69, 1003–1010.