Title: Suburban frog ponds: mapping local land cover yields different estimates of landscape composition than the National Land Cover Database

Author: Max R. Lambert

Yale University, School of Forestry and Environmental Studies

Greeley Memorial Laboratory, 370 Prospect St., New Haven, CT 06511

Email: max.lambert@yale.edu

KEYWORDS: exurban, GIS, houses, landscape ecology, land use, lawn, periurban, remote sensing, residential neighborhoods, urban ecology
Abstract – Suburban neighborhoods are rapidly spreading globally. As such, there is an increasing need to study the environmental and ecological effects of suburbanization. At large spatial extents, from county-level to global, remote sensing-derived land cover data, such as the National Land Cover Dataset (NLCD), have yielded insight into patterns of urbanization and concomitant large-scale ecological patterns in response. However, the components of suburban land cover (houses, yards, etc.) are dispersed throughout the landscape at a finer scale than the relatively coarse grain size (30m pixels) of NLCD may be able to detect. Our understanding of ecological processes in heterogeneous landscapes is reliant upon the accuracy and resolution of our measurements as well as the scale at which we measure the landscape. Analyses of ecological processes along suburban gradients are restricted by the currently available data. As ecologists are becoming increasingly interested in describing phenomena at spatial extents as small as individual households, we need higher-resolution landscape measurements. Here, I describe a simple method of translating the components of suburban landscapes into finer-grain, local land cover (LLC) data in GIS. Using both LLC and NLCD, I compare the suburban matrix surrounding ponds occupied by two different frog species. I illustrate large discrepancies in Forest, Yard, and Developed land cover estimates between LLC and NLCD, leading to markedly different interpretations of suburban landscape composition. NLCD, relative to LLC, estimates lower proportions of forest cover and higher proportions of anthropogenic land covers in general. These two land cover datasets provide surprisingly different descriptions of the suburban landscapes, potentially affecting our understanding of how organisms respond to an increasingly suburban world. LLC provides a free and detailed fine-grain depiction of the components of suburban neighborhoods and will allow ecologists to better explore heterogeneous suburban landscapes at multiple spatial scales.
Introduction

In the United States, conversion of landscapes to suburban neighborhoods is increasing at rates faster than human population growth (Liu et al. 2003, Theobald 2005). In parallel, there is growing concern about the effects of suburbanization on local ecosystem function (Falk 1976, Pouyat et al. 2009, Raciti et al. 2011, Hasenmueller and Criss 2013) and wildlife populations (Hansen et al. 2005, Skelly et al. 2010, Roe et al. 2011, van Heezik and Ludwig 2012). Rapidly-increasing suburbanization necessitates research aimed at understanding how variation in suburban landscape structure influences ecological processes (Brown et al. 2005). Large scale land cover data, such as NLCD (e.g., Homer et al. 2004, Wickham et al. 2013), has been useful for studying land use effects at large regional and global extents (e.g., Jetz et al. 2007, Bonter et al. 2010, Imhoff et al. 2010, Verburg et al. 2011, Zhang and Seto 2011, Wood et al. 2014). However, detecting developed land cover outside of intense urban centers is challenging and inaccurate (Pozzi and Small 2001, Epstein et al. 2002, Leyk et al. 2014). This is often because NLCD-type datasets are unable to easily classify pixels which have an underlying heterogeneous composition (i.e., mixed pixels; Epstein et al. 2002, Smith et al. 2002). In suburban neighborhoods, ecologically-relevant landscape components such as houses, yards, and neighborhood roads comprise areas smaller than the unit of measurement (30m pixels) for NLCD, frequently creating mixed pixels that are likely to be misclassified. Furthermore, NLCD-type data has difficulty recognizing lawn grasses (i.e., turf grass; Milesi et al. 2005, Lee and Lathrop 2006, Miller Runfola et al. 2013, Wickham et al. 2013), an important and common feature of suburban landscapes. With an estimated 9.5 million hectares of lawn grass in the United States, an area equivalent to over one third of US cropland cover (Milsei et al. 2009,) this is a problem. Because NLCD frequently misclassifies lawn grasses and mixed pixels, estimating
suburban landscape composition with these data is unreliable (Fish and Pathirana 1990, Pozzi and Small 2001, Epstein et al. 2002). Different landscape measurement tools are thus needed for suburban neighborhoods (Epstein et al. 2002, Brown et al. 2005).

The aim of landscape ecology is to understand biological patterns generated by heterogeneity in environmental features (e.g., land use; Urban et al. 1987), however our capacity to accurately estimate landscape heterogeneity is contingent upon the accuracy and scale at which we assess environmental variability (Wiens 1989). Spatial scale has been an important consideration in ecology for decades (Levin 1992, Wiens 1989, Chave 2013), with the realization that local landscape composition can be significant (Skelly et al. 2013). We are now asking questions about biological patterns at the scale of suburban neighborhoods and even households (Kaye et al. 2006, Burghardt et al. 2008, Skelly et al. 2010, Kays and Parsons 2014). Understanding landscape heterogeneity and analyzing ecosystems at multiple spatial scales is particularly important in urbanized systems (Zhang et al. 2013). It has been suggested that, to properly detect suburban landscape patterns and appropriately relate them to ecological processes, we need finer-grain land cover information than is widely-available (Brown et al. 2005).

In this study, I focus on the terrestrial uplands surrounding small freshwater ponds in suburban neighborhoods. Small ponds have been important foci for studying the influence of human land use on ecosystems as well as which spatial scales are more critical (Homan et al. 2004, Declerck et al 2006, Zanini et al. 2009). However, the reliability of land cover data at smaller scales has come into question (Trenham et al. 2004). Understanding the relative role of local landscape structures necessitates the inclusion of finer-grain data. Recent attention has focused on determining relationships between suburban landscape structure and the biology of suburban frogs (Skelly et al. 2010, Gabrielsen et al. 2013, Lambert et al. In Review). As such,
suburban ponds provide a unique opportunity to explore the validity of NLCD-type land cover data at small scales and in heterogeneous environments.

Here, I use a free, high-resolution global satellite base map in a geographic information system (GIS) to derive local land cover (LLC). While others have used high resolution imagery to create land cover maps of urban areas (Myeong et al. 2001, Akbari et al. 2003, Grove et al. 2006, Miller Runfola et al. 2013), they do so using data that are often geographically-limited and with tools that are not widely available. I compare suburban landscape composition at two scales (50m and 200m radii) around ponds of two frog species, wood frogs (*Rana sylvatica*) and green frogs (*Rana clamitans*) which vary in sensitivity to suburbanization. I explore whether LLC or NLCD estimate similar suburban landscape composition for either species and whether spatial scale is important.

**Materials and Methods**

*Study Sites and Species:* I studied a set of eight wood frog and eight green frog ponds. All sixteen ponds occur within a suburban matrix in southern Connecticut, USA and have been previously monitored for metamorphosing individuals of either wood frogs or green frogs. In addition to varying in life history strategies (e.g., explosive breeding and 2-3 month larval period for wood frogs and extended breeding with a 9-12 month larval period for green frogs) the two species vary in their susceptibility to urbanization. Wood frogs are considered a forest-dependent species that are less common in urban areas (Gibbs 1998, Gibbs et al. 2005) whereas green frogs are cosmopolitan and prevalent in urban areas (Skelly et al. 2006, 2010). Even so, wood frogs, in addition to green frogs, successfully breed in the context of human residential areas. The eight wood frog ponds range in size from 33.9-1181.3m$^2$ (mean 601.5m$^2$; SE 156.7m$^2$) and the eight green frog ponds range in size from 355.2-5159.8m$^2$ (mean 1726.0m$^2$, SE 549.0m$^2$). The number
of households surrounding these green frog ponds is substantially higher than wood frog ponds within both a 50m radius (green frog mean = 3.4 ± 0.4 houses, wood frog mean = 1.5 ± 0.4 houses, t-test P = 0.003) and a 200m radius (green frog mean = 35.5 ± 3.2 houses, wood frog mean = 15.8 ± 1.8 houses, t-test P < 0.001).

**NLCD Land Cover Analysis:** In ArcMap 10.1 (ESRI) I created 50m and 200m buffers around each wetland’s polygon and extracted recently-released 2011 United States National Land Cover Database (30m pixels) information within each buffer for each pond (Fig. 1). I then calculated the percent of the surrounding landscape of each pond, at both scales, that was occupied by Forest, Yard, and Developed land covers. I combined “Mixed Forest”, “Deciduous Forest”, “Evergreen Forest”, and “Scrub/Shrub” NLCD land cover categories into a single Forest category. My NLCD Yard category is comprised of the “Developed, Open Spaces” land cover type which predominantly represents lawn grasses and other landscaping vegetation (Wickham et al. 2013). My NLCD Developed land use category is a combination of “Developed, Low Intensity”, “Developed, Medium Intensity”, and “Developed, High Intensity” cover types. These land cover types are predominantly composed of impervious surfaces such as buildings and roads (Wickham et al. 2013).

**Local Land Cover (LLC) Derivation:** The goal of deriving LLC around suburban ponds was to create a blueprint-type schematic that represents the dominant features of residential areas (Fig. 1). I downloaded a free, high (0.3m) resolution, November 2013 global satellite basemap (Fig. 1, satellite imagery credited to ESRI World Imagery) into ArcMap. The sub-meter resolution of this map allowed for easy detection of different landscape variables. In ArcCatalog 10.1 I created new polygonal shapefiles for each pond at the 200m buffer size. When deriving land cover shapes, I created Forest, Building, Paved Surfaces, Lawn, and Garden categories. Buildings are
mostly houses but occasionally include other commercial buildings. Paved Surfaces are predominantly surface streets but also include highways, driveways, and parking lots. Garden is predominantly landscaping vegetation (not lawn grasses) but occasionally includes pools and hot tubs. In ArcMap, I edited each shapefile, creating polygons around the perimeter of each of the above-mentioned land cover types. To ensure all polygons fit together tightly, I used the snapping function. For each pond, I “dissolved” each 200m buffer by the land cover type and calculated the surface area of the five land cover categories. Finally, I “intersected” each pond’s 50m buffer with its 200m land cover shapefile and calculated the proportion of each land cover present within 50m of each pond’s perimeter.

**Statistical Analyses:** I conducted all statistical analyses in R (2.15.12, R Core Team) and set alpha to 0.05. I used simple linear regression to explore whether the proportion of LLC-derived Forest, Yard, and Developed land covers were correlated with proportions of the same land cover category derived with NLCD. I did this for each land cover category at both the 50m and 200m scales and for each species separately.

To determine whether the proportion of Forest, Yard, or Developed land covers differed between wood frog and green frog habitats using either LLC or NLCD, I used generalized linear models (GLMs) with the glm function in R for each land cover type and scale separately. I then conducted Tukey’s post-hoc tests with the glht function in the R package “multcomp” to determine pair-wise differences in land cover proportions between species and land cover datasets.

Because LLC can break down suburban landscapes into more components than NLCD, I used GLMs to compare landscape composition (forest, building, paved surface, garden, and lawn) between wood frogs and green frogs at both scales.
Results

For each pond, the relative proportions of Forest, Yard, and Developed land covers vary between LLC and NLCD data for both wood frogs and green frogs at both the 50m (Fig 2) and 200m (Supplementary material Appendix 1, Fig. A1) scales.

Regression analyses indicated few correlations between LLC and NLCD for different land covers and scales as well as for either species (Supplementary material Appendix 1, Fig. A2). For wood frogs, Forest cover was positively correlated between LLC and NLCD at the 50m scale ($R^2 = 0.46, P = 0.04$) and the 200m scale ($R^2 = 0.70, P = 0.002$). Developed land cover at the 200m scale was also positively correlated between LLC and NLCD ($R^2 = 0.95, P < 0.001$) for wood frogs. For wood frogs, there were no correlations between LLC and NLCD for Yard cover at either scale or Developed cover at 50m (all $P > 0.05$).

For green frogs, Developed cover at 200m was positively correlated between LLC and NLCD ($R^2 = 0.74, P = 0.003$). All other regressions were non-significant (all $P > 0.05$).

GLMs comparing NLCD and LLC estimates detected significant differences for each land cover at each scale (Fig. 3): Forest cover at 50m ($P < 0.001$), Yard cover at 50m ($P < 0.001$), Developed cover at 50m ($P = 0.007$), Forest cover at 200m ($P < 0.001$), Yard cover at 200m ($P < 0.001$), and Developed cover at 200m ($P < 0.001$). For Forest, Yard, and Developed covers at both the 50m and 200m scales, Tukey’s post-hoc test revealed pairwise differences between wood frogs and green frogs and between LLC and NLCD datasets (Fig. 3). Generally, NLCD estimates lower proportions of forest cover and higher proportions of anthropogenic land covers (Yard and Developed) than LLC.
At the 50m scale (Fig. 4a), LLC estimated that green frog ponds had lower forest cover (green frog mean = 49.1% ± 4.5%, wood frog mean = 75.2% ± 5.6%, P = 0.003) and higher lawn cover (green frog mean = 28.8% ± 4.2%, wood frog mean = 9.4% ± 2.7%, P = 0.002) than wood frogs. No other land cover types varied between the species at 50m (all P > 0.05). At the 200m scale (Fig. 4b), green frogs had lower forest cover (green frog mean = 37.0% ± 2.4%, wood frog mean = 67.9% ± 4.8%, P < 0.001), higher building surface area cover (green frog mean = 5.9% ± 0.6%, wood frog mean = 3.6% ± 0.6%, P = 0.013), higher garden cover (green frog mean = 5.4% ± 0.8%, wood frog mean = 3.0% ± 0.5%, P = 0.02), and higher lawn cover (green frog mean = 35.9% ± 2.2%, wood frog mean = 13.5% ± 2.5%, P < 0.001) than wood frogs.

Discussion

My results indicate that LLC and NLCD provide starkly different estimates of landscape composition surrounding suburban frog ponds. These discordant estimates are non-trivial when interpreting the upland habitat composition between wood frogs and green frogs. Ecologists are interested in understanding landscape composition at more local scales, such as the scale of residential neighborhoods. While NLCD admittedly was not designed to explore patterns at these smaller spatial extents, it provides a fundamentally different interpretation of suburban landscapes at ecologically-relevant scales than does LLC. Furthermore, LLC provides a more complex estimate of suburban landscape structure, distinguishing features like buildings from roads. LLC can also discriminate smaller patches of the landscape that are often lost in the coarser NLCD pixel size. At the scale of neighborhoods, the resolution of LLC matters for understanding ecological processes. We can have a high degree of confidence in these results.
given the data sources used here were created over a relatively short period of time (recently released 2011 NLCD data and 2013 satellite imagery).

Suburban Habitat Differences: A consistent defining characteristic between suburban green frog and wood frog ponds is yard cover. Green frog ponds are surrounded by more Yard cover than wood frog ponds, though LLC indicates a more substantial difference in magnitude between the species than does NLCD at the 50m scale.

Regardless of LLC or NLCD datasets or scale, green frogs have lower forest cover and higher yard cover surrounding ponds than do wood frogs. However, the magnitude of species-level differences varies dramatically between dataset and scale of observation. For instance, when using NLCD at the 50m scale, suburban wood frog ponds are surrounded by 55% forest cover while green frog ponds are surrounded by 5% forest cover; over a 10-fold difference. However, when using LLC at the same scale, wood frog ponds are surrounded by 75% forest cover and green frog ponds by 49% forest cover. While the difference between the species using LLC data is still significant, it is substantially less dramatic than when using NLCD datasets. Using NLCD, we would interpret green frog ponds as being nearly devoid of forest cover; with LLC green frog ponds appear more similar to wood frog ponds.

Interestingly, when looking at Developed land cover using NLCD at the 200m scale, green frog ponds appear substantially more developed than wood frog ponds With LLC, this distinction disappears and the two species have similar levels of Developed cover. While it still holds that green frogs inhabit landscapes that are more suburbanized than the landscapes inhabited by wood frogs, the more defining features of these landscapes may be relative proportions of Forest and Yard cover rather than the typically-considered Developed land covers.
**Landscaping Vegetation:** Yards, and lawns in particular, are integral features of suburban landscapes with important implications for suburban ecosystem function and species composition (Law et al. 2004, Milesi et al. 2005, Kaye et al. 2006, Burghardt et al. 2008, Pouyat et al. 2009). However, NLCD and LLC yield substantially different estimates of Yard cover with no correlations for Yard cover estimates between the two datasets. NLCD has previously been shown to be poor at recognizing lawn grasses (Milesi et al. 2005, Wickham et al. 2013), potentially leading to the incongruity between the two Yard cover estimates. At the local scale, NLCD estimated dramatically-higher Yard cover than LLC for both frogs. Given that landscaped vegetation may have important implications for aquatic pollution and amphibian populations (Lambert et al. *In Review*), accurately estimating yard cover in suburban areas is integral. Because lawns can be distinguished from other yard features in most high resolution satellite imagery, LLC provides a useful way of assessing variation in suburban lawn cover. Using LLC I was able to show that lawns, in exclusion of other yard vegetation, are dominant features surrounding suburban frog ponds and that two species breed in ponds surrounded by differing degrees of lawn cover.

**Effects of Landscape Complexity:** The uplands surrounding suburban wood frog ponds are predominantly forested. My analysis could not detect a difference in forest cover between NLCD and LLC for wood frogs, likely because these habitats are more homogeneous. Concurrently, NLCD did not detect any Developed land cover surrounding most wood frog ponds at the 50m scales whereas LLC indicated that these ponds were surrounded by substantial amounts of Developed cover. While NLCD may adequately describe large extents of forest cover in suburban landscapes, it may do so at the expense of anthropogenic land covers. Green frog suburban habitats are more complex, exhibiting a patchy, heterogeneous composition of
landscape features. Because of this, LLC and NLCD differed dramatically in forest cover estimates. Contrary to wood frog ponds, NLCD tends to estimate much higher Developed cover than LLC. Interpretations of urban landscape ecology patterns may therefore vary dramatically between LLC and NLCD data if study organisms exist in heterogeneous landscapes.

NLCD often misclassifies pixels when landscape features are smaller and land cover complexity is higher (Smith et al. 2002, 2003). It is possible that where LLC distinguishes small land cover patches, NLCD clumps multiple categories together, creating higher proportions of anthropogenic cover and lower proportions of forest cover than LLC would indicate. For some ponds, such as wood frog ponds that are predominantly one cover type, NLCD may suffice for estimating larger swaths of forest cover at within neighborhoods. However, when local landscapes are more complex, such as with green frog habitat, LLC is likely a more ideal tool.

The choice of which land cover analysis to use will depend both upon the natural history of the study organism as well as the complexity of the land use context.

Conclusions:

Wood frogs and green frogs are expected to exist at different points along a suburbanization gradient, with wood frogs at the lower end (Gibbs 1998, Gibbs et al. 2005) and green frogs at the higher end (Skelly et al. 2006, 2010). LLC data support this with forest and lawn covers being the dominant differences between the two species and garden as well as building covers playing a lesser role. Surprisingly, paved surface cover did not differ between the two species. Being able to parse apart which components of the suburban matrix (buildings, roads, lawns, etc.) drive species differences in suburban habitat could be important for ecosystem management in light of continuing suburbanization. Doing so requires tools that accurately measure landscape variables at a high resolution. Furthermore, urban gradients are important for relating ecological
phenomena to urban features (McDonnell et al. 1997, Luck and Wu 2002). The features of suburban areas are often below NLCD grain size, and so NLCD is unable to describe suburban environments in detail. Because landscape composition at local scales can have important ecological consequences (e.g., Skelly et al. 2013), using finer-grain LLC at small spatial extents will be a valuable complement to coarser-grain data (like NLCD) used over lager regions.

With suburban land cover becoming an increasingly-large component of the Earth’s surface, we need to understand how ecosystems are affected at household scales. As the availability of free, high resolution satellite imagery and GIS freeware are increasingly-prevalent, LLC data sets should allow a diversity of researchers and organizations to create detailed, accurate estimates of local suburban landscape structure and monitor environmental effects. LLC data will allow for reliable descriptions of relationships between suburban land cover and ecology. Which land cover data biologists use, as well as the scale they analyze landscapes at, has important consequences for understanding ecology, particularly in heterogeneous suburban environments.

Acknowledgments – I thank M. Holgerson, L. Freidenburg, M. Rogalski, and J. Smith for helpful discussions on these methods. D. Skelly and C. Donihue provided valuable comments on an earlier draft of this manuscript. I thank L. Freidenburg for help identifying suburban wood frog ponds and G. Giller for help targeting suburban green frog ponds and for use of his photographs.
Literature Cited


Lambert, MR. et al. In Review. Suburbanization, estrogen contamination, and sex ratio in wild amphibian populations. - PNAS.


Figure 1: Characteristic suburban ponds for wood frogs (top) and green frogs (bottom). For each species a single pond is presented (left to right) showing 2013 0.3m resolution satellite imagery with a 200m radius buffer around the two ponds, LLC-derived by translating landscape features from satellite imagery into polygons, and remotely-sensed 2011 NLCD. For LLC, dark green is forest, light green is lawn, orange is miscellaneous yard features (e.g., gardens and hot tubs), red are buildings (mainly houses), and black is paved surfaces. For NLCD, green indicates Forest, yellow represents Yard (predominantly lawn), and red is Developed cover (mostly impervious surfaces). For wood frog ponds, NLCD often does not depict some roads and low density housing around ponds. Around green frog ponds, NLCD often estimates yard cover where LLC describes forest cover. Adult frog images are courtesy of G. Giller.
Figure 2: Percent cover of Forest, Yard, and Developed covers using LLC (a,b) and NLCD (c,d) within a 50m buffer surrounding wood frog ponds (a,c) and green frog ponds (b,d). LLC and NLCD data can be compared for a given. For wood frogs, LLC describes Developed land cover that is not described by NLCD and NLCD estimates more Yard cover than LLC. For green frogs, LLC describes Forest cover where NLCD does not and NLCD more Yard and Developed cover than does LLC. White bars are Forest cover, gray bars are Yard cover, and black bars are Developed land cover. X-axis number labels represent the same pond for top and bottom panels.
Figure 3: Differences in Forest (a,b), Yard (c,d), and Developed (e,f) land cover composition in wood frog and green frog ponds at both the 50m (a,c,e) and 200m (b,d,f) scales. Black bars are LLC data and gray bars are NLCD data. Letters over bars represent Tukey’s post-hoc classifications.
Figure 4: Landscape composition breakdown using LLC data at the 50m (a) and 200m (b) scales.

Wood frog ponds are in gray and green frog ponds are hollow bars. Asterisks indicate significant differences in estimated cover between wood frog and green frog ponds for a given landscape component. In general, Green frog ponds have less forest cover but more lawn cover than wood frog ponds. Surprisingly, the landscape matrix for wood frog and green frog ponds does not differ in paved surface cover. LLC allows for a higher resolution breakdown of landscape components than NLCD.