Northeastern Regional Forest Fragmentation Assessment: Rationale, Methods, and Comparisons With Other Studies

Andrew Lister, Rachel Riemann, Tonya Lister, and Will McWilliams[1]

Abstract.—Forest fragmentation is thought to impact many biotic and abiotic processes important to ecosystem function. We assessed forest fragmentation in 13 Northeastern States to gain a greater understanding of the trends in and status of this region's forests. We reclassified and then statistically filtered and updated classified Landsat imagery from the early 1990s, and devised analysis routines that allowed for automated processing of large areas. We discuss the rationale for the study and the choices made in data set preparation and analysis routines, describe the methods used, and compare our methods with those of other coarse-scale fragmentation studies.

Introduction

The U.S. Department of Agriculture Forest Service's Northeastern Forest Inventory and Analysis unit (NE-FIA) collects data relating to quantity, quality, distribution, and health of forests from a network of ground plots distributed uniformly across 13 Northeastern States. These data are summarized and used to produce annual reports of the trends in and status of the region's forest resources. In addition to tabular summaries (e.g., McWilliams *et al.* 2002), analytical reports are produced that integrate contextual information, social data, and historical perspectives to help users interpret the numerical data. Data on fragmentation provide contextual information. For example, two counties with similar forest-area percentages can have different landscape configurations. Interpreting tabular data within the context of landscape configuration will help us gain a better understanding of the status of the forest resource and aid regional planners and decisionmakers.

Forest fragmentation also is an important issue in the ecology community. The partitioning of large, homogeneous landscape units into smaller patches by human activities and other processes influences animal behavior, plant-seed dispersal, hydrological processes, and local weather conditions (Forman 1995), all of which affect our forests. Analyzing NE-FIA forestry data through the prism of forest fragmentation can help ecologists understand regional ecological patterns.

Our objective was to design an efficient, scalable process that would produce contextual data on forest fragmentation. Specifically, we wanted to (1) provide a rationale for assessing regional forest fragmentation, (2) describe the methods used in the assessment, and (3) compare our methods with those of other coarse-scale fragmentation studies.

The protocol we developed was tailored to NE-FIA's reporting needs. Past efforts entailed manually interpreting points on a grid superimposed over aerial photography. At each point location, fragmentation metrics were recorded (Riemann and Tillman 1999). Disadvantages of this approach include high labor and materials costs and a great dependence on the quality of the photointerpreter. With the completion of the National Land Cover Data (NLCD) (Vogelmann *et al.* 2001), a 30-m Landsat-based land use/land cover classification, and the development of APACK, an efficient software application for calculating fragmentation metrics (Mladenoff and DeZonia 2001), new opportunities have emerged for measuring landscape patterns over large areas.

Before designing the procedures used in the assessment, we developed the following rationale for the analysis: to provide information for analysts and others interested in interpreting NE-FIA data with respect to patch features that are commonly reported as having a direct or indirect influence on biological systems, e.g., the average size of contiguous forest patches, their degree of isolation from other patches, shape characteristics, and length of interface between the patches and other land cover types (Forman 1995).

[1] Research Forester, Research Forester, Research Forester, and Project Leader, respectively, U.S. Department of Agriculture, Forest Service, Northeastern Research Station, FIA, 11 Campus Blvd., Ste. 200, Newtown Square, PA 19073. Phone: 610–557–4038; e-mail: alister@fs.fed.us.

We also developed a definition of forest patch that matched as closely as possible NE-FIA's definitions of "forest." For land use on an NE-FIA plot to be classified as forest, it must be at least 0.4 ha (1 acre) in extent and nearly devoid of human development (except for silvicultural treatments). For example, agricultural fields with trees or recreation areas with paths and undergrowth control would not be considered forest. On the basis of these criteria, a forest patch was defined as a contiguous area of forest cover that is at least 0.4 ha in size and differs sharply from its surroundings due to land use change, bisection by a road, or interface with a water feature such as a large river or lake. Including characteristics of forest structure in our definition might be preferable, but the data do not allow for finer distinctions beyond broad land use/land cover categories.

We are aware of only two other regional or superregional forest fragmentation assessments in the landscape ecology literature: studies by Riitters *et al.* (2002) and Heilman *et al.* (2002). Both used raster data from NLCD and devised algorithms for segmenting large images and calculating metrics. After reviewing their methods, we chose a different analytical approach, primarily because of the manner in which the NLCD data were preprocessed. Riitters *et al.* excluded roads in their analysis, and Heilman *et al.* included only major roads. We believe that the ecological effects of all road sizes are too important to ignore. Also, we wanted to correct the situation in which NLCD forest is overpredicted in areas with high tree cover but a nonforest land use, for example, a residential area with an extensive tree canopy. Because any metrics calculated depend on the accuracy of the source data set used (e.g., Riemann *et al.* 2003), we believe that this correction was critical. Finally, both Heilman *et al.* and Riitters *et al.* used subcounty-scale analysis units. NE-FIA produces statistical summaries at the county or multicounty scale, which requires different procedures than those used in the other two studies. We partitioned the landscape into political units (counties) to more closely match the reporting needs of NE-FIA.

Methods

Combining Imagery and Roads

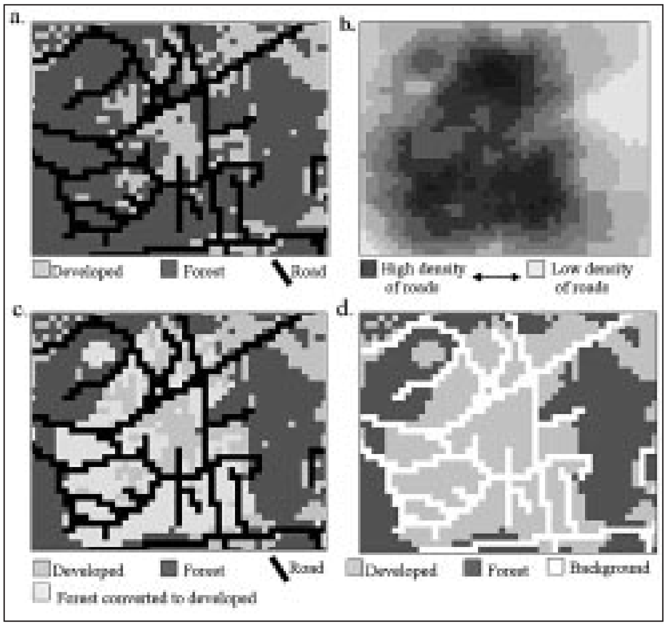
We obtained NLCD data from the U.S. Geological Survey for the 13 Northeastern States under the purview of NE-FIA (fig. 1) and then merged these data to create a contiguous, regional raster data set. We collapsed the original 21 NLCD classes into six new classes representing the land uses we were willing to consider together as a single patch (table 1) to create a new mosaic (*M*). We combined Geographic Information System (GIS) coverages of roads from the U.S. Census Bureau's TIGER/Line Files (U.S. Department of Commerce 2002) with *M* to create a new data set (*M+R*) in which each pixel of *M* that co-occurred with a road became a background or "no data" pixel in *M+R* (fig. 2a). In addition to boundaries created by roads, water and the edges of analysis units did not contribute to the edge measurements. We did encounter registration errors in various areas between *M* and the roads' data, but ignored them, assuming that the false patches created by these errors generally represent a marginal proportion of the total area and number of patches.

Figure 1.—*The States that make up the study region: Ohio, West Virginia, Maryland, Pennsylvania, Delaware, New York, New Jersey, Connecticut, Rhode Island, Massachusetts, Vermont, New Hampshire, and Maine.*

Updating and "Correcting" the NLCD Data Set for Missing Development

We had previously noted that NLCD overrepresented forest pixels in areas that include both development and high levels of

Figure 2.—*(a) roads overlaid on the NLCD image in a simplified region that includes two classes (forest and developed) (M+R); (b) results of a convolution filter that provides an index of road density (RD); (c) results of a Boolean expression that replaces forest with high road density with developed (M+R+F); (d) the final map in which patches smaller than 0.4 ha have been removed, and roads have been converted to background.*



tree canopy cover (Riemann *et al.* 2003). For these forested areas with higher road density, we applied a convolution filter (moving window) with a circular, seven-pixel-radius kernel to *M+R* so that the count of road pixels within the kernel was calculated and attached to each pixel in the output (*RD*) (fig. 2b). This output was then evaluated using Boolean logic of the form "If a pixel in *M+R* is forest, and the co-occurring pixel in *RD* has a value greater than 35, then update that pixel with the class 'developed' (table 1); otherwise, leave it with the original value of *M+R*." We decided on a threshold of pixel values greater than 35 in *RD* as indicative of high road density through a heuristic approach using different thresholds and different areas of the study region.

Approximating NE-FIA's Minimum Area Definition for Forest Land

To approximate NE-FIA's area requirements for forest classification, we eliminated all isolated patches of pixels from the map updated in the previous step (*M+R+F*) (fig. 2c) that contained fewer than four pixels of the same land cover type and replaced them with the majority land cover surrounding each updated pixel. This, in effect, defined the minimum mapping unit of *M+R+F* as 3,600 m^2 (0.9 acres) (fig. 2d).

Analysis of Reporting Units and Automation

We defined several scales of reporting unit based on the interests of NE-FIA analysts and data consumers: county, watershed, ecoregion, and State. We obtained GIS layers for county and State boundaries from the U.S. Census Bureau (U.S. Department of Commerce 2002). We designed a series of GIS-based software programs that used these GIS layers to clip *M+R+F* and process each resulting analysis unit using APACK software, as well as Environmental Systems Research Institute's ArcInfo GIS. Fragmentation metrics for each land use class from table 1 and the landscape as a whole were compiled in tabular form for each analysis unit (table 2). We do not address our choice of metrics in this article.

Table 1.—*Collapsing scheme used to convert NLCD or derived classes into our classification scheme (see text); background classes form patch boundaries but do not form patches.*

Our class	NLCD or derived class
Developed	Residential, commercial, high road density forested
Barren	Quarries, gravel, bare earth, transitional
Forest	Deciduous, conifer, mixed, woody wetlands
Natural vegetation	Shrubs, grasslands, herbaceous wetlands
Agriculture	Pasture, row crops, grains, orchards
Background	Water, roads, areas outside of the analysis region

Discussion of Methods

Riitters *et al.* (2002) did not preprocess the NLCD data other than recoding them to forest/nonforest. This approach did not meet our objectives. During our initial analyses, we determined that a single string of pixels can connect two isolated forest patches, creating a "super patch" that constitutes a large portion of the land area of the analysis unit. The methods of Riitters *et al.* (2002) are based on a sliding window and were not meant to produce patch-based measurements, whereas our requirements dictated a patch-based approach. Also, we wanted

Table 2.—*Examples of fragmentation metrics calculated for each analysis unit.*

Percent land use in	Shared edge between forest and	Other metrics
Developed	Developed	Forest edge density
Barren	Barren	Avg. corrected patch perimeter-area ratio
Forest	Natural vegetation	Avg. normalized patch area
Natural vegetation	Agriculture	Patch size summary statistics
Agriculture		Patch size histograms
		Patch connectivity metrics

to retain information on different categories of nonforest land because the nonforest land use type bordering forests is believed to affect the ecology of that forest (Forman 1995).

Heilman *et al.* (2002) preprocessed the NLCD data using a subset of the TIGER/Line roads data—U.S. interstates and routes and State and county highways—to account for the super patch problem and in recognition of the ecological impacts of roads on terrestrial ecosystems (Trombulak and Frissell 2000). They also recoded the NLCD data to forest/nonforest, and, like Riitters *et al.*, lost information by grouping all nonforest land use types in a single category.

The work of Heilman *et al.* (2002) did not meet our objectives because they omitted road classes such as rural, neighborhood, and vehicular trails, and because the boundaries of their analysis units were formed by roads rather than by political boundaries. We believed that including all available roads was important because even unpaved forest roads strongly affect the local ecology (Haskell 2000). Further, Heilman *et al.* eliminated urban areas from their analysis. We included these areas because some urban areas have significant tree cover or marginal forest (Riemann 2003).

Our methods and those of Riitters *et al.* and Heilman *et al.* (2002) share several weaknesses. First, the NLCD data have varying accuracies (Yang *et al.* 2001). By collapsing the 21 NLCD classes (table 1), we no doubt raised the overall accuracy of the data set, although measuring this directly would be difficult. At best, the NLCD forest/nonforest accuracy rates tend to range from 80–95 percent across the study region (Yang *et al.* 2001).

Second, the spatial mismatch between the TIGER/Line roads data and the NLCD image can be substantial. We experimented with several ways to address this, e.g., buffering the roads, but believed that the additional inaccuracy introduced by using the roads was offset by the ability to delineate meaningful forest patches in a way that met our definitions.

Third, the NLCD classification is driven by land cover. If an area is completely covered with tree canopy but is mowed beneath the tree canopy, e.g., in a town park, the NLCD might classify that area as forest, while the NE-FIA classification would be nonforest. This definitional mismatch is inherent in most satellite-based land cover classifications of vegetation. In our analysis, we assumed that all tree-covered areas greater than 0.4 ha are forest, although this supposition is not true. We made this assumption because no other consistently classified, national, land use/land cover maps exist at relatively fine scales. Provided that these deficiencies are recognized and understood, we believe that our method can effectively assess forest fragmentation at the regional scale.

Conclusions

One strength of our approach is the type of automation we developed. We were able to quickly and efficiently use a combination of GIS, spreadsheets, and C (programming language) programs to partition, preprocess, calculate metrics for, and compile tabular summaries of data in our analysis units. This flexibility allows us to generate metrics for any attribute of interest for which a GIS data source exists. Also, our preprocessing of the NLCD data adds an indicator of below-canopy fragmentation to areas that are tree covered on the NLCD image but replete with roads below the canopy. By including all available roads as patch-creating entities, we are in agreement with the prevailing view that any road size affects forests in numerous ways (Trombulak and Frissell 2000). Finally, by eliminating patches that do not approximate NE-FIA's forest definition, we arrive closer to the point where we can mitigate the distinction between tree cover and NE-FIA's definition of forest.

Literature Cited

Forman, R.T.T. 1995. Land mosaics: the ecology of landscapes and regions. Cambridge, MA: Cambridge University Press. 632 p.

Haskell, D.G. 2000. Effects of forest roads on macroinvertebrate soil fauna of the southern Appalachian Mountains. Conservation Biology. 14: 57–63.

Heilman, G.E.; Strittholt, J.R.; Slosser, N.C.; Dellasala, D.A. 2002. Forest fragmentation of the conterminous United States: assessing forest intactness through road density and spatial characteristics. BioScience. 52: 411–422.

McWilliams, W.H.; Alerich, C.A.; Devlin, D.A.; *et al.* 2002. Annual inventory report for Pennsylvania's forests: results from the first two years. Resour. Bull. NE-156. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 71 p.

Mladenoff, D.J; DeZonia, B. 2001. APACK 2.17 user's guide. Madison, WI: University of Wisconsin, Forest Landscape Ecology Lab. 52 p.

Riemann, R. 2003. Pilot inventory of FIA plots traditionally called "nonforest." Gen. Tech. Rep. NE-312. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 44 p.

Riemann, R.; Lister, A.; Hoppus, M.; Lister, T. 2003. Fragmentation statistics for FIA: designing an approach. In: Proceedings, 3rd annual Forest Inventory and Analysis symposium; 2001 October 17–19; Traverse City, MI. Gen. Tech. Rep. NC-230. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 146–155.

Riemann, R.; Tillman, K. 1999. FIA photointerpretation in southern New England: a tool to determine forest fragmentation and proximity to human development. Res. Pap. NE-709. Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 12 p.

Riitters, K.H.; Wickham, J.D.; O'Neill, R.V.; *et al.* 2002. Fragmentation of continental United States forests. Ecosystems. 5: 815–822.

Trombulak, S.C.; Frissell, C. 2000. Review of ecological effects of roads on terrestrial and aquatic communities. Conservation Biology. 14: 19–29.

U.S. Department of Commerce, Census Bureau. 2000. Census 2000 TIGER/Line Files [machine-readable data files]. Washington, DC: U.S. Department of Commerce, Census Bureau.

Vogelmann, J.E.; Howard, S.M.; Yang, L.; *et al.* 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering & Remote Sensing. 67(6): 650–662.

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