Generating Broad-Scale Forest Ownership Maps: A Closest-Neighbor Approach

Brett J. Butler[1]

Abstract.—A closest-neighbor method for producing a forest ownership map using remotely sensed imagery and point-based ownership information is presented for the Northeastern United States. Based on a validation data set, this method had an accuracy rate of 58 percent.

Introduction

The ownership of America's forest resources can be divided into Federal, State, and local governments; forest industry; other corporate; and family and individual ownerships. Although large variability can exist in these categories, they have proven effective in understanding how forest land is used (e.g., Haynes 2003), who receives the goods and services produced, and how private or public entities influence these trends (e.g., Sampson and DeCoster 1997). Owners are the critical link between forests and society, and a full understanding of forest resources necessitates an understanding of forest ownership patterns.

The distribution of forest owners across the country is far from uniform. Eastern forests are dominated by private owners, while western forests are dominated by public owners (Smith *et al.* 2001). The distribution of forest owners at finer scales also varies greatly due to historic land distribution policies and local economic and social forces. Most data on forest ownership have been summarized in tabular format by geographic unit, e.g., State (Smith *et al.* 2001) or county (Griffith and Widmann 2003). Although these data constitute important information, they are limited with respect to geographic resolution, do not allow for visual examination of subcounty spatial patterns, and are not conducive for combining with other spatially explicit information. A forest ownership map can overcome many of these shortcomings.

Ownerships maps (i.e., plat books) are available from county or municipal tax offices. Ideally, these maps would be accurate, publicly accessible, and in compatible digital formats. Combined with a forest map, the ownership map should depict forest ownership throughout the Nation. Unfortunately, detailed ownership records have not been assembled in electronic format for multistate regions. An exception is the Managed Area Database (McGhie 1996) that provides boundaries for major Federal and State ownerships in the United States. Other ownership maps are available at finer scales (e.g., counties) or at the State level with specific limitations, for example, only holdings greater than 200 ha (500 ac). Also, such maps are often proprietary.

Because no national maps or data sources exist for forest ownership patterns, estimation procedures are necessary. In this article, I describe a method for producing a forest ownership map for the Northeastern United States. Data from remotely sensed imagery and ground-based forest inventories are combined using a closest-neighbor approach. The accuracy of this technique is assessed using validation data, and directions for future research are discussed.

Methods

The study area for this project is the 13 Northeastern States stretching south from Maine to Delaware and west to Ohio. This area was selected because forest inventory data were readily accessible. Geographic information system layers for this study were derived from Moderate Resolution Imaging Spectroradiometer imagery (MODIS) (Hansen *et al.* 2002) and U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) plots. MODIS is satellite-based imagery. For the data used herein, the spatial resolution or pixel size is 500 m (1,640 ft) with 2000 and 2001 acquisition dates. The MODIS product used is Vegetation Continuous Fields percent forest cover data, which is produced and distributed by the Global Land Cover Facility at the University of Maryland. This information represents the percentage of each 25-ha (62-ac) pixel that is covered by tree canopies. A forest/nonforest map

[1] Research Forester, U.S. Department of Agriculture, Forest Service, Northeastern Research Station, 11 Campus Boulevard, Suite 200, Newtown Square, PA 19073. Phone: 610–557–4045; fax: 610–557–4250; e-mail: bbutler01@fs.fed.us.

was generated by assigning a value of 1 (forest land) to all pixels with a tree cover of at least 55 percent and a value of 0 (nonforest land) to all other pixels. The 55-percent minimum was selected because it generated the same regional forest percentage (61) as an independent data source (Smith *et al.* 2001).

From the FIA data, 13,686 forested plots or sample points were identified in the Northeast. The footprint of these plots covers less than 1/10 of 1 percent of the region's forest area, but the random selection procedure and large sample size ensures that the sample is representative of the broader region. For each identified plot, data on ownership of record were obtained from tax records, and each ownership was categorized as Federal, State, or local government; forest industry; other corporate; or family and individual. Forest industry includes all private holders who own a primary wood-processing facility. Other corporate includes all other businesses, associations, and tribal lands with no primary processing facilities. Family ownership includes all forest land owned by individuals or families that are not incorporated.

The coordinates of the plot locations were fuzzed to mask exact plot coordinates and prevent disclosure of landowners' identities. For each coordinate at each point, a randomly selected value of + 3,250 to – 3,250m (+/– 10,663 ft) was added.

The FIA sample points were divided into training and validation sets. One in four randomly selected sample points were reserved for the validation set. For every forested pixel on the forest/nonforest map, an ownership category was assigned based on the ownership of the closest FIA plot in the training set to generate the forest ownership map. Euclidean distances were used to find the closest plots (Environmental Systems Research Institute 2001).

For each plot in the validation set, the modeled or estimated ownership category was obtained from the forest/ownership map generated. Observed and predicted values for each validation point were used to create a confusion matrix and assess model accuracy. Although the 25 percent of plots used for validation were selected at random, the effect of this specific subset of the points on the accuracy estimate is unknown. In future efforts, we will use multiple iterations of the validation selection process to assess accuracy variability.

Results

The resulting ownership map (fig. 1) is a fair approximation of the forest ownership pattern in the Northeastern United States. The most striking feature is the vast amount of family and individual forest land in the region. As would be expected, this feature is supported by FIA tabular estimates (Smith *et al.* 2001). The map shows the major holdings by the forest industry in Maine, northern New Hampshire and northern Vermont, north-central Pennsylvania, the Adirondack region of New York, and south-central West Virginia. Large State-owned forest holdings are evident in Pennsylvania.

Of the 3,421 validation points, 58 percent were classified correctly (table 1). The most common category for a misclassified plot was family and individual ownership. This error was related to the fact that family and individual ownerships represent a plurality of forest owners in the Northeast. Part of the misclassification was due to the use of fuzzed coordinates; 26 percent of the errors were forested plots being assigned to areas classified as nonforest on the map. Other errors were attributable to the degree of accuracy and resolution of the forest/nonforest map.

Figure 1.—*This forest ownership map for the Northeastern United States was generated using a closest-neighbor approach.*

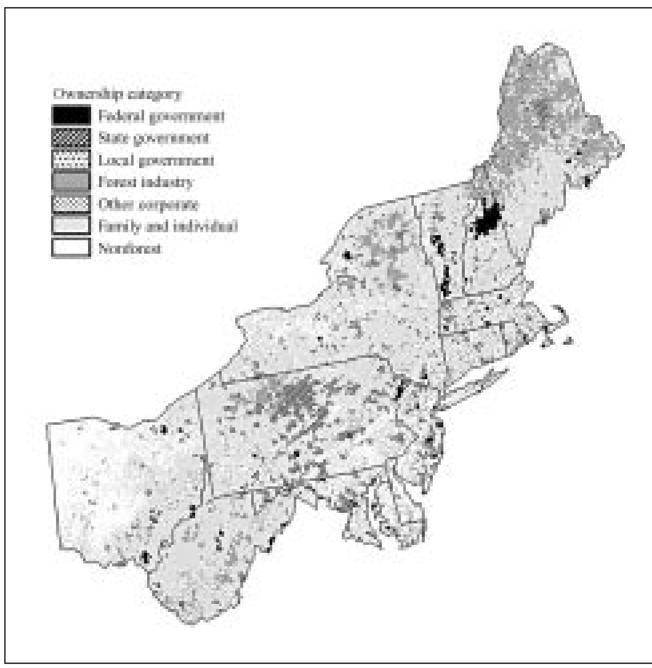

Table 1.—*Confusion matrix representing observed and predicted ownership categories based on a closest-neighbor estimation technique (numbers in parentheses represent column percentages).*

	Observed						
Predicted	**Federal**	**State**	**Local**	**Forest industry**	**Other corporate**	**Family**	**Total**
Federal	42	3	2	1	7	23	78
	(56.8)	(1.0)	(2.7)	(0.2)	(1.5)	(1.1)	(2.3)
State	0	139	3	20	26	102	290
	(0.0)	(45.1)	(4.1)	(4.9)	(5.7)	(4.9)	(8.5)
Local	1	8	6	4	11	25	55
	(1.4)	(2.6)	(8.1)	(1.0)	(2.4)	(1.2)	(1.6)
Forest industry	0	13	3	251	36	65	368
	(0.0)	(4.2)	(4.1)	(61.5)	(7.9)	(3.1)	(10.8)
Other corporate	2	21	11	40	157	204	435
	(2.7)	(6.8)	(14.9)	(9.8)	(34.6)	(9.7)	(12.7)
Family	21	111	35	80	183	1,391	1,821
	(28.4)	(36.0)	(47.3)	(19.6)	(40.3)	(66.1)	(53.2)
Nonforest	8	13	14	12	34	293	374
	(10.8)	(4.2)	(18.9)	(2.9)	(7.5)	(13.9)	(10.9)
Total	**74**	**308**	**74**	**408**	**454**	**2,103**	**3,421**

Data disclosure is a concern with the spatial display of FIA plot data. Because this method produces a relatively high rate of misclassification, data disclosure likely is not a significant issue.

Conclusions

The map resulting from the closest-neighbor approach is useful for displaying broad forest ownership patterns. This type of product would be appropriate for inclusion in a State forest inventory report or other reports concerned with forest resources across wide-ranging areas.

Our accuracy assessment, however, highlighted several underlying shortcomings. Although the general location and distribution of forest ownership may be correct, the exact locations of specific ownerships are not modeled accurately, in part because fuzzed coordinates were used to generate the map. This shortcoming limits the utility of the map for inclusion in spatial modeling projects. Thus, a percentage-based method may be more appropriate, although such a product might be inferior to a discrete map that is easier to display and interpret.

In addition to using actual coordinates, one can test other techniques for improving the accuracy of the map. The first step should be to include ancillary information, for example, data from the Managed Area Database (McGhie 1996) that depict the actual boundaries of forest owners. These data would be particularly useful for increasing accuracy in the depiction of large public ownerships. Incorporating other ancillary data and using spatial and/or nonspatial modeling techniques also could increase accuracy.

The method presented here is a relatively simple approach to generating a forest ownership map using the best available data. After additional methods are tested, this project should be expanded to include the entire Nation.

Literature Cited

Environmental Systems Research Institute, Inc. 2001. ArcInfo, version 8.3. Redlands, CA: Environmental Systems Research Institute.

Griffith, D.M.; Widmann, R.H. 2003. Forest statistics for West Virginia: 1989 and 2000. Resour. Bull. NE-157. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 119 p.

Hansen, M.; Defries, R.; Townshend, J.R.G.; *et al.* 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. Remote Sensing of the Environment. 83: 303–319.

Haynes, R.W. 2003. An analysis of the timber situation in the United States: 1952–2050. Gen. Tech. Rep. PNW-560. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 254 p.

McGhie, R.G. 1996. Creation of a comprehensive managed areas spatial database for the conterminous United States: summary project technical report. NCGIA Tech. Rep. 96-4. Santa Barbara, CA: University of California at Santa Barbara. 53 p.

Sampson, R.N.; DeCoster, L.A. 1997. Public programs for private forestry: a reader on programs and options. Washington, DC: American Forests. 95 p.

Smith, W.B.; Vissage, J.S.; Sheffield, R.M.; Darr, D.R. 2001. Forest resources of the United States, 1997. Gen. Tech. Rep. NC-219. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station. 190 p.