Quantitative classification of a historic northern Wisconsin (U.S.A.) landscape: mapping forests at regional scales

Lisa A. Schulte, David J. Mladenoff, and Erik V. Nordheim

Abstract: We developed a quantitative and replicable classification system to improve understanding of historical composition and structure within northern Wisconsin's forests. The classification system was based on statistical cluster analysis and two forest metrics, relative dominance (% basal area) and relative importance (mean of relative dominance and relative density), as computed from the original U.S. Public Land Survey (PLS) bearing-tree data. Broad forest patterns are consistent between the two metrics; yet, detailed inspection highlights different aspects of historical structure. Maps produced characterize vegetation at regional scales and reveal patterns that can be interpreted in the context of environmental constraints. Our classifications have a fairly coarse spatial grain (2.6 km^2) and fine-scale, patchily distributed ecosystems types are not represented. This resolution, however, is consistent with that of the PLS bearing-tree data, and maintaining it allowed retention of other beneficial map qualities, including quantitative representation of the data, replicability, flexibility, and an assessment of robustness and confidence. Our classifications are broadly applicable for regional-scale scientific and forest-management uses, including (*i*) assessing natural variability, (*ii*) determining the potential distribution of species, (*iii*) setting goals for ecological restoration, and (*iv*) calculating landscape change.

Résumé : Nous avons mis au point un système quantitatif et reproductible de classification pour améliorer la compréhension de la structure et de la composition passées des forêts du Nord du Wisconsin. Le système de classification est basé sur l'analyse statistique par grappes et deux paramètres forestiers, la dominance relative (% de surface terrière) et l'importance relative (moyenne de la dominance relative et de la densité relative), tel que calculés à partir des données originales sur les arbres de direction provenant de l'arpentage des terres publiques des États-Unis. Les configurations générales de la forêt sont les mêmes avec l'un ou l'autre des paramètres mais une inspection détaillée fait ressortir différents aspects de la structure passée. Les cartes qui sont produites caractérisent la végétation à l'échelle régionale et révèlent des patrons qui peuvent être interprétés dans le contexte des contraintes environnementales. Nos classifications ont une définition spatiale assez grossière (2,6 km^2) et les types d'écosystème présents à petite échelle et morcelés ne sont pas représentés. Cette résolution est cependant compatible avec celle des données sur les arbres de direction et le fait de la maintenir permet de conserver d'autres caractéristiques cartographiques bénéfiques incluant la représentation quantitative des données, la reproductibilité, la flexibilité et une évaluation de la robustesse et du degré de confiance. Nos classifications sont applicables de façon générale à des fins scientifiques et d'aménagement forestier à l'échelle locale, incluant (*i*) l'évaluation de la variabilité naturelle, (*ii*) la détermination de la distribution potentielle des essences, (*iii*) l'établissement d'objectifs pour la restauration écologique et (*iv*) la détermination des changements à l'échelle du paysage.

[Traduit par la Rédaction]

Introduction

Within the last decade, the concepts of ecosystem management and natural variability have emerged as strong frameworks for understanding and maintaining forest integrity and resilience in the face of human-induced change (Cissel et al. 1994; Thomas 1996; Wallin et al. 1996; Landres et al. 1999). Scientists and managers employing these frameworks have growing information needs. Some of these needs include precise maps of ecosystem types at multiple scales, knowledge of the proportions of ecosystem types on the landscape at different time periods, quantification of the rates of change from one ecosystem type to the another, and an understanding of the drivers of change (Wallin et al. 1996; Cleland et al. 1997; Fulé et al. 1997). Obtaining such information often requires turning to the past. Assessing natural variability itself involves acquiring data sets from multiple time periods. Many driving ecological processes are slow and (or) infrequent, requiring long-term data to characterize. For these reasons, historical data sources are receiving greater attention.

A key historical data source of interest to many ecologists and land managers is the U.S. General Land Office's origi-

Received 20 August 2001. Accepted 28 April 2002. Published on the NRC Research Press Web site at http://cjfr.nrc.ca on 4 September 2002.

L.A. Schulte,[1,2] D.J. Mladenoff, and E.V. Nordheim. Department of Forest Ecology and Management, University of Wisconsin, Madison, WI 53706, U.S.A.

[1]Corresponding author (e-mail: laschulte@fs.fed.us).
[2]Present address: USDA Forest Service, North Central Research Station, 1831 Highway 169 E, Grand Rapids, MN 55744, U.S.A.

 DOI: 10.1139/X02-082

nal Public Land Survey (PLS) records. In many regions of the United States, these data can represent vegetation cover for the period just prior to the settlement by peoples of Euro-American descent (hereafter referred to as presettlement), a period of widespread vegetation transformation. The data recorded by the General Land Office surveyors while traversing the landscape constitutes our best picture of presettlement vegetation, and because of this, the surveyors' notes have been used in numerous studies. Advantages of the PLS records as a historical vegetation data source are many and include that quantitative and spatially explicit data were collected. Because data were recorded on a grid system, they can be statistically representative, although the proper tests should be performed to confirm this (Manies et al. 2001). The PLS records also have a broad spatial coverage and a relatively fine spatial grain. Although these data have limitations (Schulte and Mladenoff 2001), research has shown that they can mirror the patterning of ecosystems present prior to settlement if the resolution of study is sufficiently coarse (Bourdo 1956; Grimm 1984; Delcourt and Delcourt 1996; Manies and Mladenoff 2000). The location and time of collection is integral to the meaning of the PLS data. The records are broadly useful in documenting the ecological conditions of a specific period or under a certain management regime (e.g., Native American, early Euro-American). For studies of natural variability, however, the PLS data are likely most useful in regions where Native American impacts were comparatively local and minimal (e.g., regions of marginal agriculture) and when they were collected prior to Euro-American settlement.

The PLS data have been used to construct presettlement vegetation maps in many midwestern states (Marschner 1974; Finley 1976; White and Mladenoff 1994; Comer et al. 1995; Delcourt and Delcourt 1996; Brown 1998; Batek et al. 1999). Development of state-level coverages has involved laborious hand transfer of presettlement data to paper or Mylar, then delineation of ecosystem boundaries using soil and topographic information. Many subjective decisions are made when following this methodology, and the result is a highly generalized coverage that is dependent on environmental data. The accuracy of such maps is difficult to assess. Quantitative mapping has been employed for smaller regions. Common approaches employ indirect ordination or geostatistical interpolation to construct models using the point-based bearing-tree data (Delcourt and Delcourt 1996; Brown 1998). These models are then applied to the entire study region. Such approaches are quantitative and replicable but also require much time to develop and can only be applied to the small, fairly homogeneous regions from which the models were derived.

Here we develop an objective method for mapping presettlement vegetation at regional scales. Our goals were to derive maps that are widely useful in further scientific and management application and improve understanding of historical vegetation types and their patterning. We achieved this goal by employing the PLS bearing-tree data in conjunction with forest structural metrics and statistical cluster analysis. Relative dominance and relative importance were chosen as structural measures as they are widely used and readily interpretable measures of tree distribution on the landscape. Cluster analysis was used for classification, because it is an objective method designed to reveal statistical patterns in data, and its use allowed us to avoid making a priori assumptions about the types of cover classes present on the historic landscape. To satisfy our requirement of an entirely quantitatively based classification, maps of presettlement vegetation were constructed at a fairly coarse resolution, with a grain or cell size of 2.6 km^2 or 1 mi^2.

Materials and methods

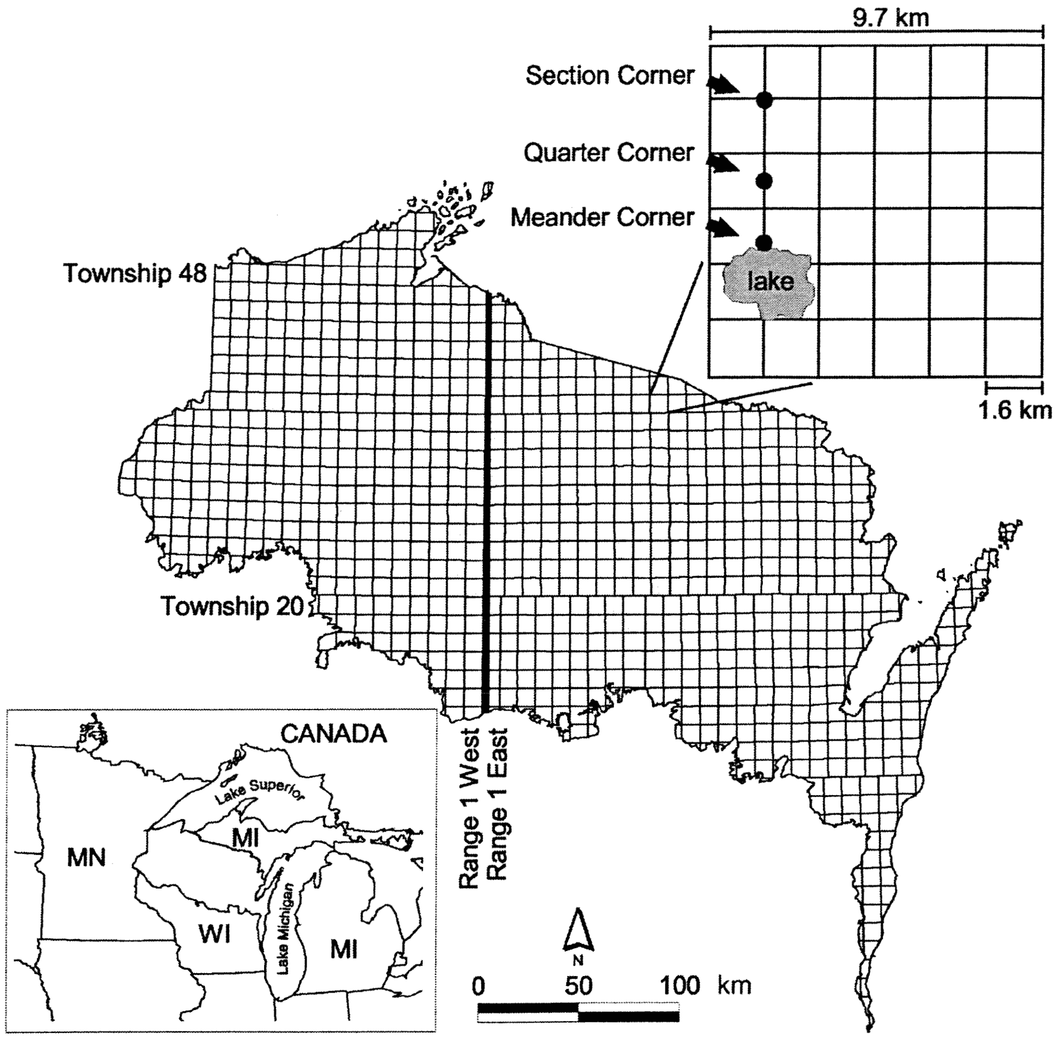
PLS records

The PLS was instituted by the federal government in 1785 and was conducted in Wisconsin between 1832 and 1866 to demarcate the land for sale, grants, and future settlement. In surveying, the state was first divided into square townships, measuring 9.7 × 9.7 km (6 × 6 mi). Townships were further subdivided into thirty-six 2.6-km^2 (1-mi^2) sections. Surveyors traversed the boundaries between all sections and, in so doing, marked the intersection of section lines (section corners) and the midpoint between section corners (quarter corners) using either a wooden post set into the ground, an earthen mound, or stones (Fig. 1; Stewart 1935). Surveyors also marked the locations where section lines crossed or circumnavigated rivers, bayous, and lakes (meander corners). At section, quarter, and meander corners, surveyors "blazed" two to four nearby trees as "bearing" trees, so that given corners could be subsequently located. Species, diameter, azimuth, and distance to the bearing tree from the corner were then recorded in notebooks.

As PLS records were not gathered for ecological purposes, they tend to be coarser than data specifically collected for vegetation mapping, and data quality can be variable (Bourdo 1956; Grimm 1984; Delcourt and Delcourt 1996; Manies et al. 2001). An area of concern surrounding use of the PLS records as a vegetation data source is the degree to which preference for or discrimination against tree species by the surveyors affects sample representativeness. Within our study area, Manies et al. (2001) have shown that elements of individual surveyor preferences did enter the equation but also that these preferences were variable, and a consistent bias across the landscape was not detected.

Again, because the PLS was not a vegetation survey, it is not always clear what the recorded information represents. One area of ambiguity important to vegetation mapping is the noting of bearing trees at the genus level or by using vague common names. Genus designations constitute a problem where the ranges of several species within the same genus overlap. For example, in Wisconsin, surveyors frequently identified bearing trees only to "birch", which could refer to either white birch (*Betula papyrifera* Marsh.) or yellow birch (*Betula alleghaniensis* Britt.) Vague common names include "spruce pine", which could have been used to denote jack pine (*Pinus banksiana* Lamb.), white pine (*Pinus strobus* L.), black spruce (*Picea mariana* (Mill.) BSP), or eastern hemlock (*Tsuga canadensis* (L.) Carrière) according to a taxonomy guide developed in the early 1900s (Britton and Brown 1913). Typical previous approaches that deal with unclear designations either (*i*) conservatively group bearing trees at the most general level (e.g., *Betula*

Fig. 1. Public Land Survey Township and Range divided the land into townships, measuring 93.2 km^2. Townships were further divided into 2.6-km^2 units, called sections (right inset), and survey data were collected along section boundaries. Bearing trees were marked and recorded at section, quarter, and meander corners.



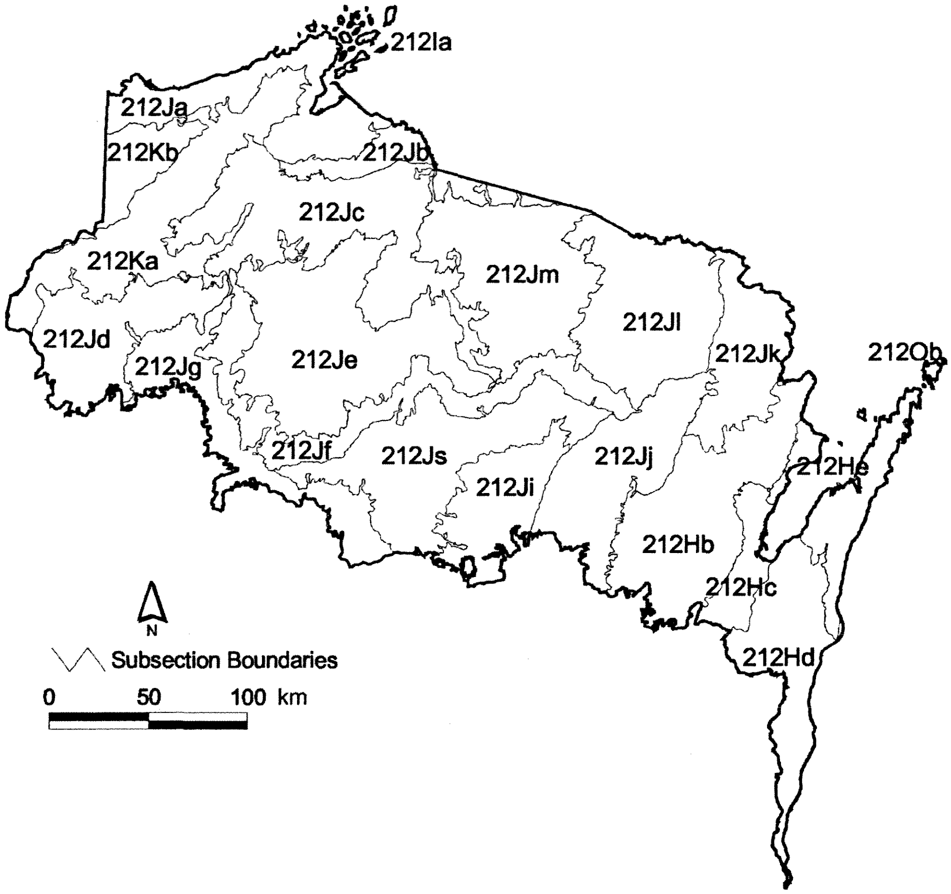
spp.; Grimm 1984; Schulte and Barnes 1996; Radeloff et al. 1999) and (or) (*ii*) subjectively divide ambiguous bearing trees into separate species groups based on phytosociological associations with soil attributes and other tree species (Finley 1976; Grimm 1984; Barrett et al. 1995; Zhang et al. 2000). We did not opt for either of these approaches. Grouping species to the most general level constitutes a loss of ecological information and their inclusion created awkward classes in preliminary phases of analysis. Previous uses of phytosociological associations have been very subjective and are undesirable for precise and quantitative mapping. Instead, we employed an objective methodology developed by Mladenoff et al. (2002), which uses logistic regression to differentiate generic and ambiguous bearing trees to the species level. Criteria used in the regression included bearing-tree diameter, mean tree density, associated tree species, and subsection and land type association (LTA) from the U.S. Forest Service's ecoregion classification (Keys et al. 1995; WDNR 1999). In the current data set, of the total number of bearing trees within a genus (e.g., jack, white, and red (*Pinus resinosa* Ait.) pine are all included in the pine genus), 10.4% of pine trees, 32.4% of maple trees, 80.4% of birch trees, 37.2% of oak trees, and 41.1% of ash trees were differentiated using this technique. Unclear designations of spruce and elm were not differentiated here. Spruce was not differentiated because the number of unambiguous designations was too few. Elm remained generic because the unambiguous data were poorly distributed across the study area.

Study area

The study area is the Wisconsin portion of the U.S. Forest Service's Vegetation Province 212 (Fig. 2; Keys et al. 1995; WDNR 1999). This region covers 78 000 km^2 or, roughly, the northern half of Wisconsin and contains about 800 townships, 30 000 sections, 90 000 corners, and more than 200 000 bearing-tree records.

The climate within this region of Wisconsin has been described as humid continental. It is characterized by extreme

Fig. 2. U.S. Forest Service Province 212 in Wisconsin (Keys et al. 1995; WDNR 1999). Interior lines are the boundaries of subsection-level ecoregions.

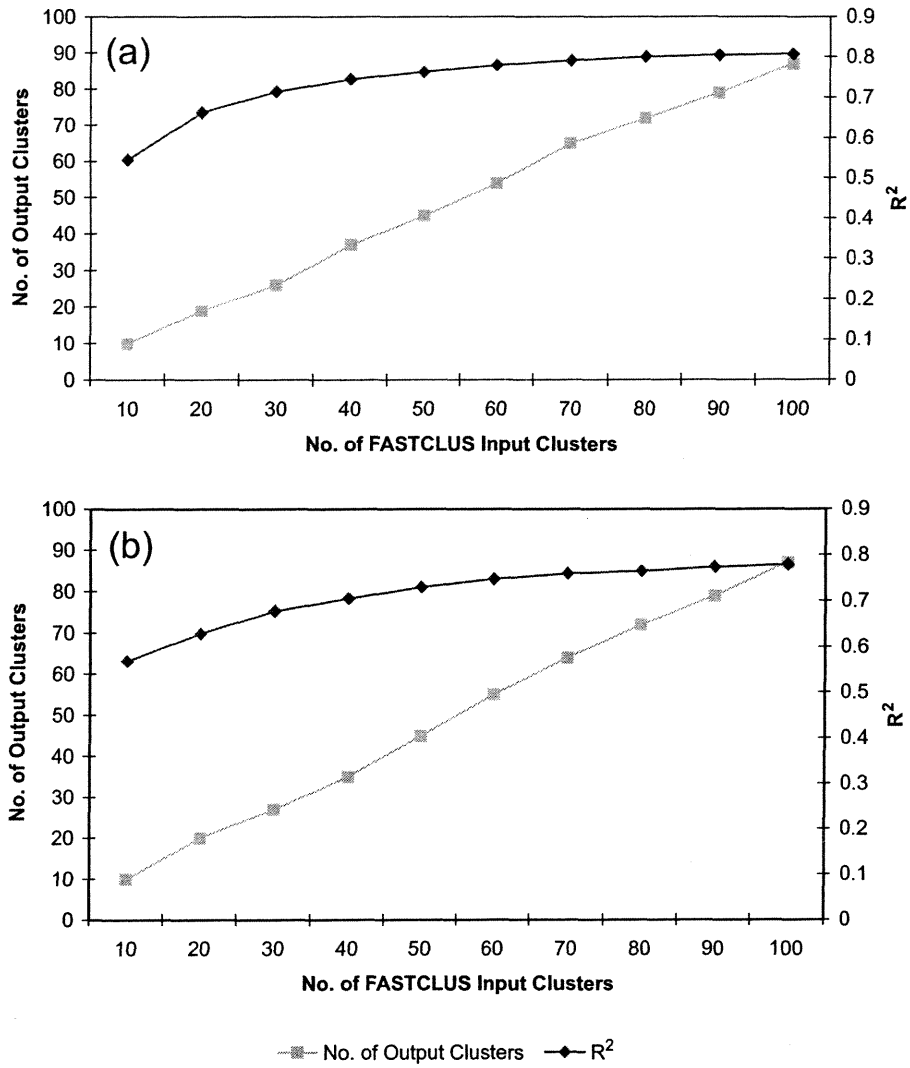


variations in temperature and moderate amounts of precipitation (76 and 87 cm), which is well distributed throughout the year. Mean January temperature is –12°C; however, temperatures of less than –34.4°C are reported from northern Wisconsin weather stations nearly every winter. Mean July temperature is 20°C. Mean length of the growing season varies between 80 days in the northeast to over 160 days on the Door Peninsula and in Manitowoc and Sheboygan counties along Lake Michigan (Karl et al. 1990).

Major plant communities of this region, as described by Curtis (1959), include conifer–hardwood, pine, and boreal forest. Hemlock – northern hardwoods, a division of conifer–hardwood forest, is the predominant land cover type in this region and is most commonly found on silty or loamy soils. Currently dominant species include sugar maple (*Acer saccharum* Marsh.) and linden (*Tilia americana* L.), although eastern hemlock and yellow birch were also common prior to widespread logging and forest clearing in the late 1800s (Curtis 1959; Mladenoff and Pastor 1993). Pine ecosystems are found in regions with droughty soils prone to fire, with the most extensive region of pine located in the upper northwestern corner (Curtis 1959; Radeloff et al. 1999). Boreal forest is located adjacent to Lake Superior on either side of the Bayfield Peninsula. Although described by Curtis (1959) as boreal, this forest type may not be typical of that found more extensively in nearby Ontario, Canada. Pockets of forested wetland, characterized by soils that are very moist to inundated for some portion of the year are located throughout the upland matrix and common wetland trees include tamarack (*Larix laricina* (Du Roi) K. Koch), black spruce, northern white-cedar (*Thuja occidentalis* L.), balsam fir (*Abies balsamifera* (L.) Mill.), and black ash (*Fraxinus nigra* Marsh.). In addition to the ecosystems mentioned by Curtis (1959), Mladenoff and Pastor (1993) also cite aspen – white birch and pin cherry (*Prunus pensylvanica* L.f.) as major ecosystems within the western Great Lakes region. Minor ecosystems within the Wisconsin portion of Province 212 include alder thickets, open bogs, fens, sedge meadows, bracken-grasslands, and beaches (Curtis 1959).

Prior to Euro-American settlement, the vegetative composition of northern Wisconsin was largely determined by physical (e.g., climate, geology, soils, etc.) and biological (e.g., competition, dispersal, herbivory, etc.) factors, and the interactions between them (e.g., disturbance regimes). Anthropogenic effects on vegetation were likely minimal and localized when compared with southern Wisconsin and to postsettlement. Low levels of charcoal and ragweed pollen within lake sediments (Webb 1974; Swain 1978; Clark and Royall 1996) and a low number of major archaeological sites (Gartner 1997) suggest smaller Native American popu-

Fig. 3. R^2 versus initial FASTCLUS cluster number for (*a*) relative dominance and (*b*) relative importance data.



lations in northern Wisconsin, a conclusion that is consistent with other areas of marginal agriculture (Cronon 1983; Vale 1998). Of physical and biological factors, postglacial species migration played a key role in determining species presence and relative abundance. Following the retreat of the Wisconsin glaciation, primary tree colonization occurred by spruce, tamarack, fir, and aspen between 10 000 and 13 000 years BP (Davis 1981). Approximately 10 500 to 11 000 years BP, oak and pine migrated into the region. Eastern hemlock and American beech are relatively recent arrivals, with establishment occurring approximately 3200 and 6000 years BP, respectively (Davis et al. 1986, 1994). Because of slow migration rates, it is possible that hemlock and American beech (*Fagus grandifolia* Ehrh.) may not have completed their westward expansion at the time of the PLS (Davis 1984).

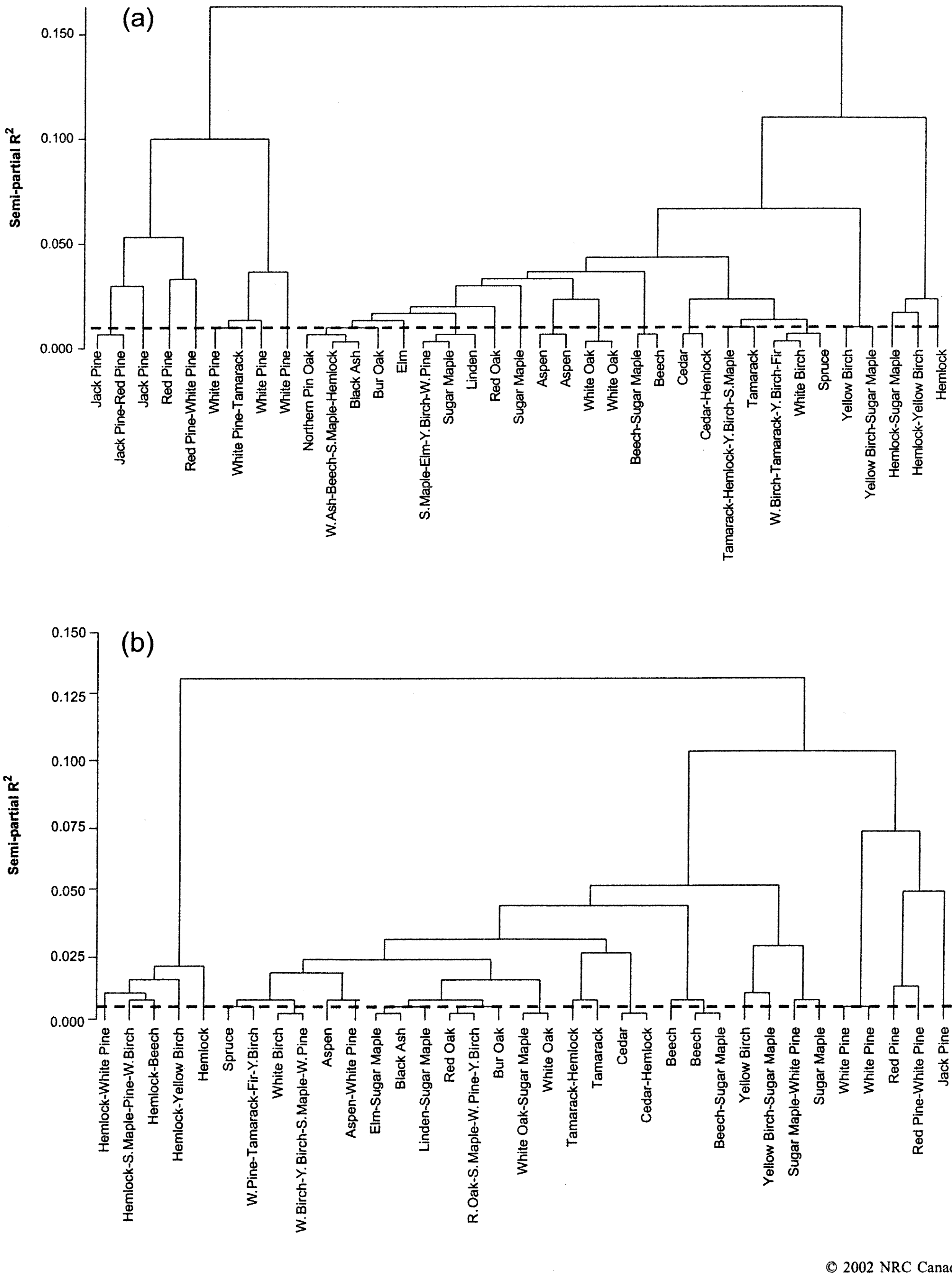
Between the mid-1800s and early 1900s, this region was subject to heavy logging followed by intense slash fires (Fries 1951). The combination of these disturbances favored early successional species including quaking aspen (*Populus tremuloides* Michx.), paper birch, and pin cherry, which subsequently became the dominant cover types (Stone and Thorne 1961; Spencer and Thorne 1972; White and Mladenoff 1994). Attempts at farming also followed initial forest clearing, but these efforts were largely abandoned during the 1920s and 1930s (Fries 1951; Gough 1997). Much of the land reverted to public ownership in the 1930s and 1940s, and today forestry and recreation are the predominant land uses.

Data processing

Structural measures

Classification of PLS corner and quarter-corner data into forest types began with a data aggregation program developed by He et al. (2000). This program, MWINDOWS, was used to combine all bearing-tree data within a section (1.6 × 1.6 km or 1 × 1 mi cell) and to compute relative dominance and relative importance for each tree species in the section. Relative, rather than absolute, measures were used because of data quality issues associated with the PLS survey, such as the preference for trees within mid-diameter ranges (25.4–40.6 cm; Bourdo 1956). In cases where a local sample

Fig. 4. Dendrogram produced through the hierarchical agglomerative clustering procedure on (*a*) relative dominance and (*b*) relative importance data. Broken line shows level of pruning according to the 1% rule (each class must contain at least 1% of total cells).



deviates from broader averages, a relative estimate reduces the magnitude of error. Because of the incorporation of relative measures, however, we were unable to distinguish regions of overall high or low biomass or importance.

Dominance, a commonly used forest metric, sums the basal area over all trees of a species within a cell, and uses DBH (diameter of the tree at breast height; 1.4 m) measurement taken from the PLS records. Relative dominance expresses basal area of a species in comparison with all other species found on the site (Cottam and Curtis 1956). Relative dominance of species i within a cell, or RDom_C_i, is calculated according to the following:

$$[1] \qquad \text{RDom\_}C_i = \left(\frac{\sum_{j=1}^{m} \text{basal area}_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m} \text{basal area}_{ij}} \right) \times 100$$

where n is the number of species and m is the number of individuals of a given species in the cell. Basal area$_{ij}$ is the basal area of individual j of species i. High measures of relative dominance suggest that a large proportion of the biomass at a site is attributable to a single species.

The version of relative importance, or RIV_C_i, used by MWINDOWS averages percent density and percent basal area of a given tree species in relation to all other tree species within the cell according to

$$[2] \qquad \text{RIV\_}C_i = \frac{\text{RDen\_}C_i + \text{RDom\_}C_i}{2}$$

where the relative density of a species within a cell, RDen_C_i, is calculated as

$$[3] \qquad \text{RDen\_}C_i = \left(\frac{\sum_{j=1}^{m} C_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m} C_{ij}} \right) \times 100$$

where n and m are the same as described above and C_{ij} is individual j of species i. High values of relative density mean that the majority of the stems within the cell are attributable to a single species. Although the relative importance metric provides a more holistic view of vegetation structure at a site (Curtis and McIntosh 1951), the meaning of values with moderate magnitude can be unclear. For example, a relative importance value of 50% for a species could mean that the species has moderate values for both dominance and density or that one of the forest metrics has high values while the other has low values. A classification based solely on relative density was not attempted, because when used alone, density does not provide a measure of the biomass allocation.

The MWINDOWS program outputs the values of relative dominance and relative importance for each species for each PLS survey section within the study area. This output was then exported to Statistical Analysis Software (SAS) for cluster analysis (SAS Institute Inc. 1990).

Clustering

We used statistical clustering to group the PLS data into forest types and made no a priori assumptions regarding the types of ecosystems that could be found on the landscape. The large amount of data (30 352 cells) to be classified required a two-step approach. In step one, values of relative dominance for the 27 most common bearing-tree species were first fed into the FASTCLUS procedure, which was developed to handle large data sets (SAS Institute Inc. 1990). FASTCLUS uses an iterative procedure and a minimum distance rule to assign samples to clusters and is highly successful at finding clusters inherent within data sets. FASTCLUS was performed multiple times using a variable number of starting clusters seeds in beginning the iterative procedure. The number of starting clusters ranged between 10 and 100, in intervals of 10. The number of output clusters (clusters remaining after the iterative procedure was finished) and associated R^2 values were then graphed to determine which cluster level should be used for further analysis (Fig. 3). R^2, a measure of the variation explained by the clustering procedure, continually increases with increasing numbers of input clusters; however, our objective was data simplification, and in order for the classification to be generally useful, we wanted to explain as much of the variance as possible in the fewest clusters. We thus chose a cutoff of <0.03 increase in R^2, and FASTCLUS output derived from using 40 initial cluster seeds was used in the second phase of classification.

The second phase of classification involved hierarchical agglomerative clustering, to observe the relationship between the output clusters from the first stage. We used Ward's minimum-variance method (Ward 1963) as a hierarchical agglomerative clustering procedure, as it is generally accepted to perform well (Everitt 1980; Milligan and Cooper 1987). In preliminary analyses using the PLS data, this method also performed better than other clustering algorithms according to subjective ecological criteria (close proximity of similar ecological groups within the dendrogram). This process was repeated for the relative importance data.

For further data reduction and generalization, we further reduced the number of classes by eliminating clusters that were infrequently represented on the landscape (<1% of total cells). To accomplish this, we pruned the dendrogram produced by the hierarchical agglomerative clustering to the level of combination resulting from the 1% rule (Fig. 4). Ecologically, this means that rare groups were eliminated from the classification. Final classes were named according to the following conventions: (*i*) single-species classes were defined as clusters in which the highest ranking species represented at least 30% of the relative dominance or relative importance and was at least two times as dominant or important as the second ranking species; (*ii*) double-species classes were defined as clusters that did not meet the criteria for a single species class but for which the sum of the relative dominance or relative importance of the two highest ranking species was greater than 50%, and the third ranking species was less than half as dominant or important as either of the first two; and (*iii*) the remainder of the classes were defined as mixed and named according to the four top-ranking species.

Table 1. Land cover classes distinguished by cluster analysis on (A) relative dominance and (B) relative importance.

(A) Relative dominance.

Class	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Whole set
Aspen	3.6	3.5	3.4	2.3	3.6	3.6	3.6	3.5	3.5	3.4	3.5
Black ash											
Black ash – white birch – sugar maple – white pine			7.1								
Beech	4.3	3.6	3.6	3.8	3.1	1.1	1.3	3.5	3.5	3.7	3.7
Beech – sugar maple						3.2					
Beech – hemlock – sugar maple – cedar							3.0				
Cedar	1.7	3.3	2.6	1.6	1.5	3.8	1.7	3.4	3.9	1.3	4.1
Elm	1.8	1.8	1.6	1.9	1.8	1.7	2.0	1.9	1.6	1.9	1.5
Hemlock	14.8	7.3	16.9	7.4	7.8	15.1	14.4	14.1	15.1	8.0	7.5
Hemlock–beech					1.3						
Hemlock–cedar	3.0	3.2		2.7	2.8		2.5			2.9	
Hemlock – sugar maple		4.7		5.7	5.0					5.2	5.2
Hemlock – white pine		4.0		4.2	4.1					4.3	
Hemlock – yellow birch	7.5	7.0	7.2	6.9	6.7	7.2	5.4	8.1	7.4	6.8	8.6
Jack pine	1.4	1.3	1.4	2.4	1.5	1.4	2.3	1.4	2.4	2.4	1.4
Jack pine – red pine	2.0	2.7	2.5	1.4	2.0	2.5	1.6	1.5	1.5		2.5
Jack pine – white pine								1.5			
Linden											
Linden – sugar maple – yellow birch – elm	3.6	4.4	2.6	1.9	1.8	1.5	1.9	3.4	2.8	1.9	5.3
Northern pin oak											
Northern pin oak – bur oak – black ash – sugar maple	2.3	1.8		1.9	3.2	2.6	2.6	1.8		2.0	2.0
Red oak	1.9	1.5	1.3	2.0	1.5	2.0	1.4	1.9	1.4	1.9	1.4
Red pine	4.2	3.3	3.8	5.4	4.0	3.7	3.5	3.8	3.8	3.4	3.6
Red pine – white pine	4.1	3.7	4.0	3.2	3.8	4.0		3.9	4.0	4.6	3.9
Spruce						1.0	1.0			1.6	
Spruce – tamarack – white pine – white birch	2.7	3.3		2.2	3.7			3.3	2.3		3.1
Sugar maple	3.9	3.2	3.8	4.3	4.6	4.0	5.8	4.6	3.7	4.5	3.9
Sugar maple – white pine – yellow birch – white birch		4.6									
Tamarack	1.4	1.9	3.8	4.3	2.0	1.7	4.2	1.8	1.6	1.7	1.9
Tamarack – white pine		2.2		2.6		2.5		2.7		2.3	2.4
Tamarack – hemlock – yellow birch – sugar maple		2.9	2.8	2.9	3.0	3.0	3.0	2.9		3.0	2.9
Tamarack – white pine – aspen – white birch									2.6		
Tamarack – white pine – yellow birch – spruce	4.8										
White birch					1.1	2.0	2.7			2.0	
White oak	2.8	2.2	2.7	2.7	2.7	2.8	3.0	2.8	2.9	2.9	2.7
White pine	17.5	12.9	18.4	12.0	13.2	14.1	16.9	16.2	16.6	14.5	18.0
White pine – yellow birch – sugar maple – hemlock				4.6	4.4	4.2				4.1	
White pine – red pine – jack pine – aspen							4.6				
White pine – sugar maple – black ash – bur oak									4.6		
Yellow birch	5.6	4.9	6.0	5.0	5.0	6.2	7.1	6.3	6.2	5.0	4.9
Yellow birch – sugar maple	5.2	4.9	4.7	4.5	4.8	4.9	4.6	5.7	5.0	4.7	6.2
Yellow birch – tamarack – fir – hemlock									3.6		
No. of classes	22	26	20	26	27	25	24	23	22	26	23

Table 1 (*continued*).

(B) Relative importance.

Class	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Whole set
Aspen	1.0	1.7	3.8	4.9	2.6	3.8	3.7	1.4	3.6	3.5	1.4
Aspen – white pine	3.2										2.9
Aspen – white pine – red pine – white birch								2.6			
Beech	1.5	5.0	1.5	5.6	1.6	1.3	1.4	1.1	1.4	1.4	1.4
Beech – sugar maple	3.2		3.5		3.3	3.4	3.3	3.4	3.3	4.1	3.3
Bur oak											
Bur oak – red oak – sugar maple – aspen	1.7		2.6	2.1		1.9		1.6	1.6	2.1	2.0
Bur oak – red oak – sugar maple – elm		6.5									
Bur oak – red oak – sugar maple – white pine							1.8				
Bur oak – red oak – white pine – aspen					1.7						
Cedar	1.6	4.1	4.4	2.9	4.4	4.2	3.3	4.6	3.1	4.4	4.5
Hemlock	6.9	12.0	10.5	6.7	6.3	6.4	11.6	10.4	6.3	10.7	6.9
Hemlock–beech	1.8		1.9		1.9	1.9	1.7	1.9	2.3		1.8
Hemlock–cedar	3.0						2.8				
Hemlock – cedar – sugar maple – yellow birch									2.9		
Hemlock – sugar maple	5.9	5.3	5.8	5.9	5.3	5.7		4.9		5.5	5.9
Hemlock–tamarack	3.6					3.4	3.1			3.3	
Hemlock – tamarack – yellow birch – cedar				6.4							
Hemlock – white pine		4.0	3.9	4.0	4.9	4.1	3.9	4.1			4.2
Hemlock – yellow birch	6.7	6.6	6.7	7.2	6.8	6.6	6.6	6.3	7.7	6.7	7.0
Jack pine	1.7	3.0	2.9	3.0	3.0	2.9	2.9	2.9	1.6	1.7	2.9
Jack pine – red pine	2.5								2.4	2.4	
Linden											
Linden – sugar maple – yellow birch – elm	2.3			5.4						2.7	
Northern pin oak											
Northern pin oak – black ash – beech – sugar maple				1.6							
Red pine	3.1	3.6	5.7	3.5	3.4	3.2	3.2	3.7	4.9	2.6	3.8
Red pine – white pine	3.8	3.7	3.0	4.4	4.5	4.6	5.8	3.7	3.2	4.4	4.0
Spruce	1.4	1.5	1.6	1.4	2.2	1.5	1.4	1.5	1.1	1.3	
Spruce – tamarack – white pine – white birch											3.8
Sugar maple	4.2	7.1	5.5	4.3	9.1	7.6	4.0	4.0	3.7	4.4	3.9
Sugar maple – elm – black ash – linden			4.4				5.6	3.9		3.1	4.8
Sugar maple – elm – black ash – yellow birch					2.1	4.8					
Sugar maple – northern pin oak – elm – black ash	3.1								5.3		
Sugar maple – white pine								3.9	3.4		3.2
Sugar maple – white pine – yellow birch – white birch			4.7				3.8	3.7			
Tamarack	5.2	3.2	2.5	5.6	1.9	2.4	2.8	2.0	2.2	3.2	2.7
Tamarack – hemlock – yellow birch – sugar maple			4.0		3.9			3.5	3.7		3.8
Tamarack – white pine					3.2			3.2	3.0		
White birch	1.8	5.2	2.4	1.5	2.8	1.6	3.0		2.3		2.6

Table 1 (*concluded*).

(B) Relative importance.

Class	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Whole set
White birch – white pine – aspen – hemlock										4.4	
White oak	2.6	2.8	1.9	2.4	2.7	2.6	1.9	2.8	1.8	1.8	2.4
White pine	15.5	11.1	10.0	12.3	13.1	12.1	9.7	9.5	11.9	13.7	11.0
Yellow birch	4.1	4.0	6.8	4.1	4.1	4.3	4.4	4.7	5.3	4.1	4.7
Yellow birch – sugar maple	4.9	5.0		4.9	5.3	5.0	4.8	4.8	10.6	5.4	5.3
Yellow birch – tamarack	3.6						3.5			3.1	
Yellow birch – tamarack – fir – spruce						4.8			1.2		
Yellow birch – tamarack – fir – hemlock		4.5									
No. of classes	27	20	23	22	24	24	25	26	26	24	25

Classification robustness and confidence

We assessed (*i*) the statistical robustness of our classifications with data-dividing, cross-validation techniques and (*ii*) the confidence that a mapped cell belonged to a particular class using a distance measure. These techniques were used in lieu of a true accuracy assessment; an assessment of true map accuracy could not be obtained because quantitative vegetation data separate from the PLS records do not exist for the period.

Cross validation was performed using data-dividing techniques and discriminant analysis, a classification technique independent of cluster analysis. Whereas cluster analysis is designed to delineate groups within data sets for which no previous categories exist, discriminant analysis assigns unclassified data into previously defined categories (Digby and Kempton 1987; McGarigal et al. 2000). In cross validation, we first divided our data into 10 random segments of approximately equal size. We next derived classifications using 90% of the data in cluster analysis and used discriminant analysis to classify the remaining 10% of the data. This step was performed 10 times, eliminating each 10% segment of the data in turn. Comparisons could then be made between how the 10% test data were classified (*i*) in the original cluster-based classification using the whole data set and (*ii*) by discriminant analysis. Such cross-validation techniques provide a measure of classification robustness and an understanding of the causes of misclassification.

We calculated the distance measure as an indication of the confidence in the membership of a cell to its mapped cover class. Distance measures are frequently used in quantifying the similarity between a cell and the class to which it is assigned and are helpful in identifying outliers (Manly 1994; Lowell and Jaton 1999). An advantage of such measures is that they can be mapped spatially, to show regions of high and low confidence. Our measure uses the value of either relative dominance or relative importance, respectively, for each of the 27 tree species common within the study area on a cell basis. In this case, a cell is an individual PLS section. The "Euclidean distance" between an individual cell value and the mean value of the cluster to which the cell belongs was calculated. For example, the mean value of relative dominance for cluster j of species i (M_{ij}) was calculated as

$$[4] \qquad M_{ij} = \frac{\sum_{q=1}^{k} \mathrm{RDom\_}C_{iq}}{k}$$

where $\mathrm{RDom\_}C_{iq}$ is the relative dominance value for species i within cell q and k is the number of cells in the cluster. The distance, D_{jq}, between the cell q and the center of cluster j was calculated as

$$[5] \qquad D_{jq} = \left(\sum_{i=1}^{n} \left| M_{ij}^{2} - \mathrm{RDom\_}C_{iq}^{2} \right| \right)^{-0.5}$$

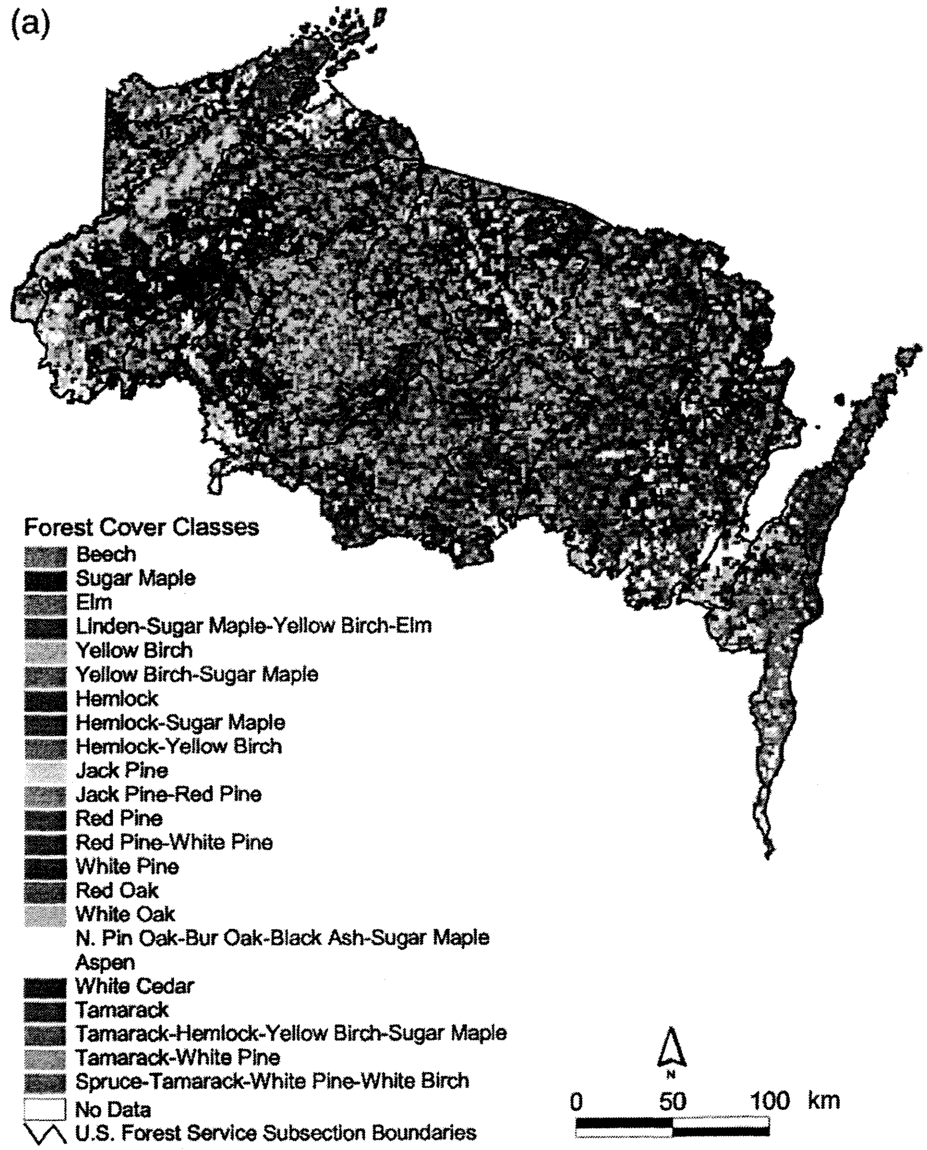
where n is the number of species found within the cluster. The value of D_{jq} for each cell, which is unitless, was then mapped. Greater distances equate to less confidence regarding membership of a cell to its mapped forest cover class.

Results

Cluster analysis

Twenty-three and 25 classes were described by cluster analysis on relative dominance and relative importance, respectively (Table 1, see Whole set column). Although the number of classes found significant using relative importance was slightly higher, it is still within range of the number of classes differentiated in cross validation of the relative dominance classification (20–27 classes; Table 1A). Single species classes of highly dominant species produce strong clusters regardless of metric used (Table 1); beech, sugar maple, yellow birch, hemlock, jack pine, red pine, white pine, white oak (*Quercus alba* L.), aspen, white-cedar, and tamarack classes were delineated using relative dominance and using relative importance. Mixed classes were also very robust to structural measure as hemlock – yellow birch, red pine – white pine, spruce – tamarack – white pine – white birch, yellow birch – sugar maple classes were common to both classifications. Classes of species that achieve large diameters played a slightly greater role in the relative dominance classification (e.g., red oak (*Quercus rubra* L.); jack pine – red pine) and classes of species that are frequently present in high densities were defined by the relative impor-

Fig. 5. Presettlement forest classification based on (*a*) relative dominance and (*b*) relative importance.

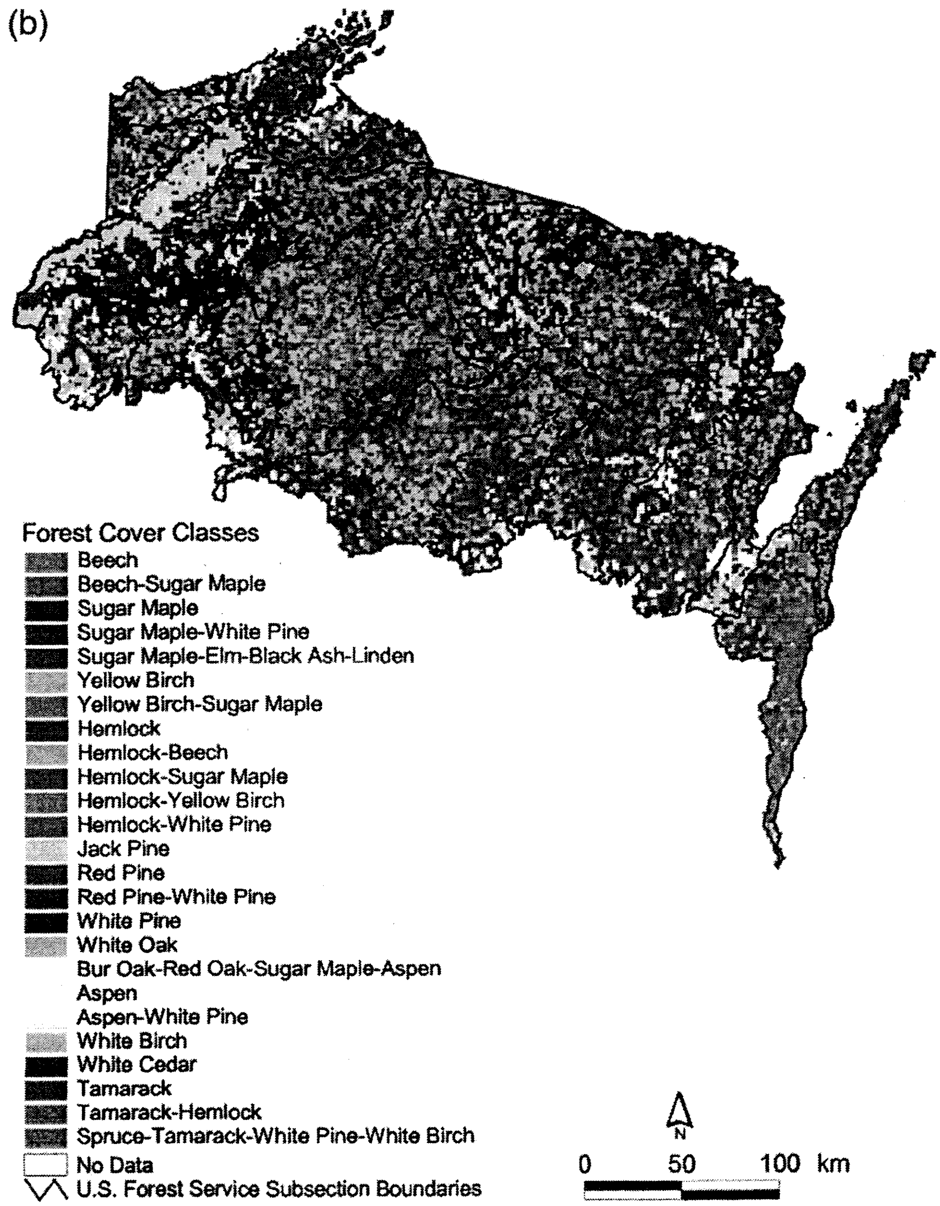


tance classification (e.g., white birch), although these trends are less striking. Generally, relative importance required more groups to describe the variation in the hemlock – northern hardwoods cover type. Beech – sugar maple, hemlock–beech, hemlock – white pine, and sugar maple – white pine were all found to be significant using relative importance, but absent from the relative dominance classification.

The spatial overlap of classes using the two structural measures is strong and similar vegetation patterns are visible in maps based on either of the structural measures (Fig. 5). Pine classes (jack pine, red pine, white pine, red pine – white pine) dominate on the sand plains within U.S. Forest Service subsections 212Jk, 212Jm, and 212Ka (Table 2, Fig. 2). Hemlock – northern hardwoods classes (hemlock, hemlock – sugar maple, hemlock – yellow birch, sugar maple, yellow birch) are found on the moraines and till plains in subsections 212Jc, 212Je, 212Js, and 212Jj in the central portion of the state. Classes that contain American beech and northern white-cedar are present in the eastern portion of the state, along Lake Michigan and Green Bay (subsections 212Hd and 212He). Swamp conifers (tamarack, tamarack – white pine, tamarack – hemlock, spruce – tamarack – white pine – white birch) dominate the Lake Superior lowlands in the northeast corner of Wisconsin (subsections 212Ja and 212Kb) and are patchily located on the pitted outwash plains (subsections 212Jc, 212Jm, and 212Jl). Oak classes (red oak, white oak, northern pin oak (*Quercus ellipsoidales* E.J. Hill) – bur oak (*Quercus macrocarpa* Michx.) – black ash – sugar maple, bur oak – red oak – sugar maple – aspen) are sporadically distributed along the southern boundary of Province 212.

The visual overlap observed in the maps holds up numerically in calculation of percent agreement between the classifications (Table 3). Although classes given the same name

Fig. 5 (*concluded*).



(e.g., "red pine" within the relative dominance classification and "red pine" within the relative importance classification) cannot be directly equated because of differences in the structure of the data from which they were derived, percent agreement between similarly named classes is generally high (≈70%). When percent agreement is not maximized between two groups with the same name, it is usually because a finer level of detail was distinguished by one of the structural measures and not the other. For example, percent agreement between aspen categories using the two structural measures is low at 36.2%; however, another 45.3% of the cells distinguished as aspen using dominance were differentiated into the double species class of aspen – white pine using relative importance (Table 3). This later aspen – white pine class was not found to be significant using relative dominance as a metric and, thus, was merged into the single species aspen class. American beech poses a similar situation. The overlap between beech classes is low at 37.8%; however, another 50.4% of cells with high values of dominance for American beech are found within the mixed class of beech – sugar maple as differentiated by relative importance (Table 3).

Classification robustness and confidence

Robustness in relative dominance classification

An assessment of the robustness of the clustering procedure, as performed by cross validation, shows that the classifications describe the statistical properties of the data well.

Table 2. Landform and soils characteristics of subsections within the U.S. Forest Service's hierarchical ecoregion classification system (Albert 1995; Keys et al. 1995; WDNR 2000).

Subsection	Landform	Soils
212Hb: West Green Bay Till Plain	Undulating till plain underlain by limestone dolomite bedrock	Rocky, podzolized sandy loams interspersed with peat and muck
212Hc: Green Bay Clayey and Silty Lake Plain	Flat plain of lacustrine clays	Leached red calcareous clays
212Hd: Manitowoc Till Plain	Flat to rolling till plain of clayey ground and end moraines	Leached red calcareous clays
212He: Door and Escanaba peninsulas and Lake Plain	Various landforms of lacustrine origin underlain by limestone and dolomite	Diverse sandy to loamy soils, calcareous and poorly drained where bedrock near surface
212Ia: Lake Superior and islands	Glacial lake plain and ice scour	Leached calcareous silt and clay, peat extensive in wetlands
212Ja: Lake Superior Clay Plain	Glacial lake plain and water-reworked moraine	Leached calcareous red loams and clays; peat extensive in wetlands
212Jb: Gogebic-Penokee Iron Range	Thin glacial till over volcanic bedrock ridge	Red loams and sandy loams
212Jc: Winegar Moraines	Ice stagnated features with steep ridges, kettle lakes, and wetlands	Acidic, rocky, red sandy loam or loamy sand with fragipans common, peat extensive in wetlands
212Jd: St. Croix Moraine	Stagnation and end moraine	Neutral sandy loam, small inclusions of peat
212Je: Central/Northwest Wisconsin Loess Plains	Undulating, nearly level plain	Acidic silt loams, poorly drained and covered with wind-blown silt
212Jf: Perkinstown End Moraine	Hummocky end moraines and outwash channels	Acidic sandy loam, small inclusions of peat
212Jg: Lincoln Formation Till Plain, mixed hardwoods	Dissected glacial till plain	Poorly drained silt loams over acidic loamy till
212Ji: Rib Mountain Rolling Ridges	Steeply dissected till plain with exposures of granitic bedrock	Well-drained shallow silt loams
212Jj: Green Bay Lobe Stagnation Moraine	Parallel end moraines separated by outwash channels and till plain	Neutral sandy loam to loam, small inclusions of peat
212Jk: Spread Eagle/Dunbar Barrens	Bedrock knobs surrounded by extensive outwash plains, some pitted	Droughty outwash sands
212Jl: Brule and Paint rivers, drumlinized ground moraine	Gently rolling drumlins and adjacent swampy depressions	Well-drained sandy-loamy or loamy sand, may have silt cap of loess
212Jm: Northern Highlands Pitted Outwash	Pitted outwash plain with coarse-textured moraines	Acidic sands
212Js: Lincoln Formation Till Plain, hemlock–hardwoods	Dissected glacial till plain	Poorly drained silt loams over acidic loamy till
212Ka: Bayfield Sand Plains	Pitted outwash plain: flat terraces intersected by hummock sediments	Deep loamy sands low in organic matter
212Kb: Mille Lacs Uplands	Lake Superior lobe ground moraine and drumlin fields	Acidic loams
212Ob: Green Bay and islands	Glacial lake plain and ice scour	Calcareous loams, poorly drained where bedrock near surface

In sequentially clustering 90% of the relative dominance data, 16 of the 23 classes (70%) were reproduced all 10 times and 20 of the 23 classes (87%) were reproduced at least 8 times (Table 1A). The hemlock – sugar maple class was represented the lowest number of times in the cross-validation data sets (4), and was still differentiated by the whole data classification. The percentage of cells that were placed into this class was robust (4.7–5.7%), as is generally true for the percentage of cells falling within the whole data set classes (Table 1A). Cases where deviation within a row is large can usually be attributed to achievement of a finer level of classification within a cross-validation data set. For example, the percentage of cells within the beech class varies between 1.1 and 4.3%. Closer inspection, however, shows that the two times when the percentage of cells within beech classes is low (1.1 and 1.3%), a secondary American beech class also occurs: beech – sugar maple in set 6 and beech – hemlock – sugar maple – cedar in set 7, respectively (Table 1A). When the percentages of these classes are added to the more general beech class, the range in percent cells of beech drops to between 3.1 and 4.3%.

Results of the discriminant analysis corroborate the robustness of the relative dominance classification described above and demonstrate where misclassification occurs (Table 4A). Percent agreement for similarly named classes is high (>70%), and agreements of 100% are found for several classes including elm, jack pine, red oak, and tamarack. Low percent agreements are found for hemlock – yellow birch

(54.1%), linden – sugar maple – yellow birch – elm (55.4%), and white pine (57.5%). In such cases, a large percentage of the cells are classified into a finer level class that does not occur in the whole data set classification. For example, 24% of the hemlock – yellow birch cells were classified as hemlock – white pine by discriminant analysis, but this hemlock – white pine class does not occur in the classification based on the whole data set. Of all classes, linden – sugar maple – yellow birch – elm is the least robust, as determined by consistently low percent agreement. Even in this case, however, the majority of the cells misclassified by the discriminant analysis are placed within closely related classes (e.g., 12.6% yellow birch – sugar maple, 9.7% elm, 4.6% sugar maple).

Robustness in relative importance classification

Assessment of the relative importance based classification shows similar trends to classification on relative dominance, while also showing that this structural measure produces more idiosyncratic classes. In repeatedly clustering 90% of the data, 13 of the 25 classes (52%) were reproduced all 10 times and 17 of the 25 classes (68%) were reproduced at least 8 times (Table 1B). These values are much lower than those for relative dominance (70 and 87%, respectively). Furthermore, the whole data set classes of aspen – white pine and sugar maple – white pine were differentiated by the cross-validation data sets only one and two times, respectively. Although a spruce class was differentiated by all cross-validation data sets, cells with high relative importance values for spruce were lumped into a mixed class of spruce – tamarack – white pine – white birch when the whole data set was analyzed. A similar number of cells were originally designated spruce within the whole data set; however, this class was lumped with a white pine – tamarack – fir – yellow birch as its semipartial R^2 value fell below the level of pruning established by the 1% rule.

Results of the discriminant analysis reveal that the erratic whole data set aspen – white pine class is a finer level designation of the aspen and red pine – white pine classes present in all of the cross-validation data sets (Table 4B). In the discriminant analysis, 64.9 and 17.0% of the aspen – white pine cells were classified into the aspen and red pine – white pine classes, respectively. The whole data set sugar maple – white pine class was formed from an amalgamation of the cross-validation data set classes of sugar maple (37.0%), sugar maple – elm – black ash – linden (20.7%), and white pine (19.6%). The singly occurring spruce – tamarack – white pine – white birch class contains predominantly spruce cells (38.3%) but also contains a substantial percentage of cells that were classified as pure tamarack (19.2%), white pine (9.2%), and white-cedar (7.5%) as differentiated by the discriminant analysis. With the exception of these three idiosyncratic whole data set classes, percent agreement between relative importance classes with the same name is very high (>70%), and the beech and jack pine classes can obtain 100% agreement.

Membership confidence

In the calculation of Euclidean distance from cell value to cluster mean, large distances were interpreted as poor classification performance. We interpreted our overall clustering as robust as few cells are contained within the greatest distance category (52.0–119.1; Fig. 6), and the overall difference between clustering on two different metrics was small. For relative dominance, the distance to cell was 27.2 ± 9.1 (mean ± SD). Distance to cell based on relative importance was 25.3 ± 8.0.

Although our methodology generally performed well for most cells, the classification system employed here performed poorly for a portion of the cells located on the edge of the study area (Fig. 6). These cells usually did not contain a full complement of data (8 data points, 16–24 bearing trees) and, therefore, were more difficult for the clustering procedure to correctly classify. Both relative dominance and relative importance classifications also performed poorly for several cells within subsection 212Jg. This region corresponds to cells for which the values of bur oak were high. Using both relative dominance and relative importance, single species bur oak clusters were distinguished during the primary phase of cluster analysis, using the FASTCLUS procedure, and were later eliminated using the 1% pruning rule. This rule was likely too strict for bur oak, which formed tight single-species clusters within the first phase of classification. Merging bur oak cells with the next-nearest cluster resulted in straggly clusters with large measures for Euclidean distance. In comparison with relative importance, classification on the relative dominance metric performed poorly within subsection 212Ka, the Bayfield Sand Plains. This is a region of nutrient poor, sandy soil that was prone to frequent burning prior to Euro-American settlement and supported dense stands of young jack pine trees (Radeloff et al. 1999). Because tree diameters were small, overall basal area within these stands was frequently low. Incorporation of density information allows the relative importance metric to produce tighter clusters and to better capture the character of these regions.

Discussion

Our goals were to (*i*) develop an objective method for classifying presettlement vegetation, (*ii*) derive maps widely useful to scientific and management applications, and (*iii*) improve understanding of historical vegetation types and their patterning. Our successes are discussed in context of the classification approach, map explanation and application, and ecological interpretation.

Classification approach

Because the PLS provides the most complete data representing presettlement land cover across much of the United States, these records have been frequently employed in studies of historical vegetation. A frequent goal in vegetation mapping using the PLS records is a high-resolution reconstruction (Finley 1951; Marschner 1974; Comer et al. 1995; Delcourt and Delcourt 1996; Brown 1998), which involves tradeoffs in which other desirable factors regarding classification are sacrificed (e.g., map independence, map accuracy). Rather than achieving a high resolution, our primary goals were (*i*) a quantitative and easily replicable classification, (*ii*) an independent classification, and (*iii*) a flexible classification. A fourth goal was to provide measures of robustness and confidence to associate with map products.

Table 3. Percent agreement matrix between relative dominance and relative importance whole data set classifications.

	Relative dominance									
Relative importance	Aspen	Beech	Cedar	Elm*	Hemlock	Hemlock – s. maple	Hemlock – y. birch	J. pine	J. pine – r. pine*	Linden – s. maple – y. birch – elm*
Aspen	36.2		0.1	0.7	0.0				0.7	
Aspen – white pine*	45.3		0.3			0.1	0.1		4.4	0.6
Beech		**37.8**	0.2							
Beech – sugar maple*	0.4	50.4	1.9	5.7		0.6				4.9
Bur oak – red oak – sugar maple – aspen*	1.5					0.2	0.0		0.8	1.4
Cedar	0.6	0.7	**80.6**	0.7	1.3	2.5	1.9		0.5	2.1
Hemlock	0.1		0.7	0.5	**76.8**	0.9	12.1			
Hemlock–beech*	0.1	9.9	8.1	0.7	6.4	4.4	2.8			0.3
Hemlock – sugar maple	0.4		1.1	2.5	9.7	**73.9**	10.0			2.3
Hemlock – white pine*	0.5		1.8	0.9	2.8	2.4	14.4		0.1	0.4
Hemlock – yellow birch	0.3		0.5	1.6	1.4	1.6	**53.3**			0.2
Jack pine								**100**	58.9	
Red pine			0.1			0.1			23.4	
Red pine – white pine	0.1			0.2					3.5	0.1
Spruce – tamarack – white pine – white birch	3.9		2.2	0.5	0.1	0.1	0.4		1.8	1.6
Sugar maple			0.2	0.7		5.7				11.2
Sugar maple – elm – black ash – linden*	1.1	0.5	0.7	81.1		1.4	0.3		0.5	**37.2**
Sugar maple – white pine*	1.3		0.1			0.6				15.9
Tamarack	1.9	0.1	0.4		0.0				2.6	0.3
Tamarack – hemlock – yellow birch – sugar maple	0.3	0.4	0.4	0.7	1.3	1.5	3.6		0.3	1.8
White birch*	4.3	0.2	0.4	0.7	0.1	1.5	0.9		1.4	4.9
White oak	0.8	0.1		0.5					0.7	0.6
White pine	0.7		0.2	0.2					0.4	0.1
Yellow birch	0.2			0.2			0.1			0.2
Yellow birch – sugar maple	0.2			2.1		2.6	0.2			14.2

Note: Columns sum to 100. Bold values indicate overlap between groups with the same name.
*Class unique to each respective classification.

Quantitative and replicable

Many previous attempts at mapping presettlement vegetation have involved qualitative use of the PLS records (Marschner 1974; Finley 1976; Comer et al. 1995). Production of these coverages involves making many subjective decisions and are extremely time intensive. For example, Finley (1951, 1976), in his *Original Vegetation of Wisconsin*, employed subjective decision-making in three regards: (*i*) in determining the species of bearing trees given ambiguous common names or recorded generically by the Government Land Office surveyors, (*ii*) in determining ecosystem types, and (*iii*) in determining the location of boundaries between adjacent ecosystem types. Rules can be developed so that these decisions are consistently applied throughout the mapping process; however, broad-scale qualitative mapping efforts are difficult to repeat. Because of the lengthy time required for qualitative mapping, replication is not frequently attempted, and the amount of confidence that can be placed in a map is never assessed.

In contrast, the technique we develop used quantitative decision-making and statistical differentiation of ambiguously defined bearing trees, forest cover classes, and the boundaries between adjacent forest cover classes. Classifications portrayed are replicable; they can be derived independently using rules provided. Quantitative classification requires weeks as opposed to months; thus, maps can be quickly updated as new information or techniques become available or if an alternative objective requires that different rules be employed.

Independence

Finer mapping resolution of presettlement vegetation is often achieved through qualitative incorporation of additional data or through detailed quantitative modeling. Marshner (1974), Finley (1976), and Comer (1995) qualitatively use the PLS bearing-tree data in combination with survey plat maps and topographic maps to distinguish fine-scale features (e.g., wetland, brush). The resultant maps, however, may not be used as independent layers in analysis and must bear the label of "potential" vegetation. Quantitative modeling approaches have included ordination and interpolation of the bearing-tree data (White and Mladenoff 1994; Delcourt and Delcourt 1996; Brown 1998; Batek et al. 1999; Manies and Mladenoff 2000). For example, Delcourt and Delcourt (1996) draw on detrended correspondence analysis (DCA) in combination with survey plat maps to delineate vegetation

Pin oak – bur oak – b. ash – s. maple*	R. oak*	R. pine	R. pine – w. pine	Spruce – tamarack – w. pine – w. birch	S. maple	Tamarack	Tamarack – hemlock – y. birch – s. maple	Tamarack – w. pine*	W. oak	W. pine	Y. birch	Y. birch – s. maple
0.7	0.2	0.1	0.1		0.3	0.4			2.1	0.0		
1.7	1.0	2.2	7.5	3.1		0.5	0.1	5.8	0.5	2.6		0.1
1.0									0.5	0.0	0.1	
14.9	2.9	0.1	0.2	0.1	3.9			0.6	4.5	1.7	0.1	0.6
21.1	76.1	0.5	0.9	0.4	0.1			0.3	2.6	1.3		0.1
2.3	1.0	0.1	0.2	1.3	0.6	0.4	5.8	2.2		0.6	0.5	2.6
0.2				0.1								
2.0	0.7		0.1				0.8	0.3	0.1	0.3		0.2
0.7			0.1	0.1	1.0		0.6		0.4	0.1		2.4
0.2	0.2		1.4	0.3			0.6	4.0		12.7		0.1
0.2	0.2			0.1			2.1		0.1	0.4	28.3	9.8
		0.1	0.3						0.1	0.2		
0.2		**83.9**	3.2	0.6		0.2						
0.2		9.4	**76.6**	0.1				1.2	0.1	2.6		
0.7	0.5	0.6	2.5	**49.0**	0.1	2.6	6.4	36.2		2.7	0.3	2.1
0.5	1.9	0.1		0.2	**74.8**	0.2			1.3	0.0		0.3
41.0	4.8	0.1		0.3	0.6		0.6	1.0	6.1	1.1	0.1	1.4
1.8	0.7	0.3	1.7	0.7	7.4		0.1	1.8	2.2	9.7		0.1
1.5	1.0	0.6	0.6	2.5		**91.2**	4.8	18.7	0.2	0.1		
1.7	0.7	0.6	0.2	3.6	0.3	4.4	**76.2**	12.7	0.2	0.2	3.1	3.2
1.8	1.9	1.0	1.7	36.9	0.5		0.6	1.2	0.2	1.4	0.2	4.6
3.6	3.1	0.3	0.1		0.1		0.1	0.1	**76.6**	0.6		0.1
0.8			2.8	0.2				13.8	0.5	**58.1**	0.1	0.1
0.2	0.7					0.2	0.3			1.6	**65.2**	18.6
1.3	2.4	0.1		0.1	10.4		0.9		1.6	2.0	2.0	**53.8**

boundaries. Plat maps, however, were hand drawn by 19th-century government draftsmen who worked from the surveyor notes and never visited the survey site; substantial subjective interpolation of vegetation boundaries must have occurred. Using the plat maps introduces substantial uncertainty, and despite this, Delcourt and Delcourt (1996) fail to consistently reduce the mapping resolution (from 2.6 km^2 to between 0.7 and 2.6 km^2). Brown (1998) employs indicator kriging in conjunction with fuzzy-c means clustering to derive forest types. This interpolation technique, however, is based on the presence or absence of a tree species at a given survey point, which may not be justifiable given the structure of the PLS data (two to four trees recorded per corner). With so few trees recorded, confidence that a species not mentioned at a given corner was absent from the surrounding forest is low. To overcome this problem, Brown (1998) incorporates soil wetness data and sacrifices the independence of the map.

We also had as one of our objectives a quantitatively based vegetation coverage independent of topographical or soils information. Although such information was not directly incorporated, the goal of independence was not completely reached; environmental characteristics were indirectly incorporated into the differentiation of ambiguous bearing trees using the Mladenoff et al. (2002) technique. Their methodology uses a variety of data, including subsections and LTAs from the U.S. Forest Service's ecoregion classification (Keys et al. 1995; WDNR 1999), in logistic regression on a survey corner basis. Although no topographical or soils information was directly included into the methodology, the U.S. Forest Service's subsection- and LTA-level ecoregions are based on aspects of climate, physiography, and soils. For some classes (e.g., hemlock, beech), no ambiguity existed in their dominant species and the differentiation technique was not applied; these classes should not be limited in application. Within the pine classes (e.g., jack pine, red pine – white pine), only a small proportion of bearing trees (10.4%) were differentiated using this technique; these classes should be applicable for most uses. Birch, on the other hand, was recorded generically by the surveyors in most instances; 80.4% of all birch bearing trees required differentiation. In this case, incorporation of environmental data into the Mladenoff, et al. (2002) technique, albeit indirectly, can exclude our maps from some ap-

Table 4. Results of discriminant analysis for relative dominance on (A) set 2 and (B) set 4.

(A) Set 2.

	Whole set									
Class	Aspen	Beech	Cedar	Elm	Hemlock	Hemlock – s. maple	Hemlock – y. birch	J. pine	J. pine – r. pine	Linden – s. maple – y. birch – elm
Aspen	**88.0**	1.7								2.3
Beech		**89.7**					0.4			
Cedar			**63.4**							
Elm				**100**			0.4			9.7
Hemlock					**87.6**	0.6	7.4			
Hemlock–cedar*		2.6	32.8		8.5	6.4	8.1			
Hemlock – s. maple						**80.1**	2.6			4.6
Hemlock – w. pine*			2.2		3.8	5.1	24.1			
Hemlock – y. birch						2.6	**54.1**			
J. pine								**100**	1.4	
J. pine – r. pine									**91.5**	
Linden – s. maple – y. birch – elm		1.7				3.2				**55.4**
Pin oak – bur oak – b. ash – s. maple		2.6								3.4
R. oak	1.0					0.6	0.4		1.4	
R. pine									1.4	
R. pine – white pine	1.0								4.2	
Spruce – tamarack – w. pine – w. birch	6.0		0.7			0.6	0.7			1.1
S. maple		0.9								4.6
S. maple – w. pine – y. birch – w. birch*		0.9								2.3
Tamarack										
Tamarack – hemlock – y. birch – s. maple							1.5			3.4
Tamarack – w. pine	3.0									0.6
W. oak	1.0						0.4			
W. pine										
Y. birch										
Y. birch – s. maple			0.7			0.6				12.6

plications, such as quantitative modeling against environmental characteristics.

Flexibility

Two forest structural measures were included to provide a more comprehensive and flexible classification system. We suggest using the two classifications in combination, as each provides a slightly different view of presettlement forest structure. For example, species that achieve large diameters (e.g., red pine, elm, and red oak) are more prominent in the classification on relative dominance, while species that occur in dense stands (e.g., white birch) are highlighted in the relative importance classification. Our classification system maintains further flexibility in the development of hierarchical relationships between land cover types. Groups within the dendrograms resulting from the polythetic agglomerative clustering procedure can be either lumped or further differentiated depending on the level of detail desired (Fig. 4).

Measures of robustness and confidence

Because natural resources scientists and decision makers require knowledge regarding the certainty associated with the data sources they are using (Lunetta et al. 1991; Lowell and Jaton 1999), another objective was to obtain a measure of the robustness and confidence of our maps. Our cross validation suggests that both single-species classes (e.g., beech, white oak) and mixed-species classes (e.g., red pine – white pine, hemlock – yellow birch) can be highly robust. Spatial overlap between relative dominance and relative importance was high (Table 3); however, the relative dominance metric generally exhibited stronger performance with the data dividing techniques. The spatial measure of confidence also highlights regions where classification performance may be lower (Fig. 6). Regions of lower confidence include a series of cells on the edges of the study area, where few bearing-tree data existed to incorporate into the aggregation algorithm, and for cells in which our 1% pruning rule was too strong to capture their character (e.g., bur oak; Fig. 4).

Map explanation and applications

The forest classification system developed here uses the point-based PLS bearing-tree data to provide a continuous description of regional scale presettlement land cover. We began with no a priori conception of the types of ecosystems

Pin oak – bur oak – b. ash – s. maple	R. oak	R. pine	R. pine – w. pine	Spruce – tamarack – w. pine – w. birch	S. maple	Tamarack	Tamarack – hemlock – y. birch – s. maple	Tamarack – w. pine	W. oak	W. pine	Y. birch	Y. birch – s. maple
1.7										0.2		
										0.2		
				2.4	1.8		2.2			1.3		
1.7									1.1	0.4		1.9
5.2							2.2	1.3				
					0.9							
								1.3		9.6		
3.4										0.2	26.6	9.9
										0.4		
8.6				1.2	0.9				3.3	0.4		1.2
75.9					1.8			1.3	1.1	0.4		1.2
1.7	**100**		0.9		0.9			1.3	5.4	0.9		1.2
		91.4	3.4									
		8.6	**93.1**	2.4			3.4			0.4		0.6
			1.7	**85.7**			3.4	2.6		1.3		1.2
					80.7							
					0.9				1.1	24.5		
						100						
				6.0			**85.4**	1.3			2.5	5.6
1.7				2.4			3.4	**91.0**		1.7		
			0.9		0.9				**88.0**	0.4		
										57.5		
										0.4	**70.9**	16.1
					11.0					0.2		**60.9**

present at the regional scale and allowed our procedure to define classes based on natural patterns in the data. Using the same philosophy and gradient analysis, Curtis (1959) defined major patch-level ecosystem types within Wisconsin. A common first division using such an approach is to delineate upland cover types from lowland cover types; Brown and Curtis (1952) ordinated upland communities and Christensen et al. (1959) followed suit with lowland communities. The limitations in the resolution of the PLS data precluded us from making such a division and from attempting a traditional community-level characterization. Manies and Mladenoff (2000), through reconstruction of the PLS, showed that the coarse distribution of data points is not capable of representing fine-scale, community-level patterning when the data are used by themselves.

Given that our classification spans such upland–lowland boundaries, and traditional patch-level ecosystems in general, how should it be used and interpreted? Classes output from the clustering algorithm represent trends in tree species dominance or importance over regional scales and tree species associations at the 2.6-km^2 resolution. Despite the broad extent and coarse resolution characterized, we were able to capture the major plant communities defined by Curtis (1959) in his community-level approach, including conifer–hardwood, pine, and boreal ecosystems. Our classification further discriminates forested wetland types, where they form the dominant cover type. Ecosystem types that are only found in small patches (e.g., brush) are not represented and are more difficult to study using the PLS data (Schulte and Mladenoff 2001).

Our classifications are broadly applicable for forest science and management uses including (*i*) assessing the range of natural variability, (*ii*) determining the potential distribution of species prior to Euro-American settlement, (*iii*) pinpointing regions and setting goals for ecological restoration, and (*iv*) calculating landscape change. For many applications, the classification based on relative dominance may be optimal, because values are more clearly interpretable and this classification was found to be more robust. However, if a presettlement coverage is required as a baseline to compare against modern satellite imagery based coverages, we recommend using the classification based on relative importance. Relative importance is more comparable with information obtained via spectral signals of the forest canopy;

Table 4 (*concluded*).

(B) Set 4.

	Whole set											
	Aspen	Aspen – w. pine*	Beech	Beech – s. maple*	Bur oak – r. oak – s. maple – aspen	Cedar	Hemlock	Hemlock– beech	Hemlock – s. maple	Hemlock – w. pine	Hemlock – y. birch	J. pine
Aspen	**91.4**	64.9							0.6			
Beech			**100**	78.1				53.7				
Bur oak – r. oak – s. maple – aspen		2.1			**86.0**	0.7			0.6			
Cedar						**74.2**						
Hemlock						0.7	**96.6**	22.2	8.4	3.0	0.5	
Hemlock – s. maple				1.9		11.3	1.0	20.4	**84.8**	2.2	0.5	
Hemlock – tamarack – y. birch – cedar*												
Hemlock – w. pine				1.0		6.0		3.7	0.6	**88.1**		
Hemlock – y. birch						2.0	2.4		1.7	4.5	**97.1**	
Jack pine												**98.7**
Linden – s. maple – y. birch – elm*				4.8	5.3				1.1		0.5	
Pin oak – b. ash – beech – s. maple*		2.1		10.5	5.3	1.3			0.6			
R. pine												
R. pine – w. pine		17.0				0.7						1.3
Spruce*		2.1				0.7			0.6	0.7		
S. maple				2.9		2.0			0.6			
Tamarack		7.4										
W. birch		1.1							0.6	0.7		
W. oak	8.6				3.5							
W. pine		3.2		1.0								
Y. birch											1.0	
Y. birch – s. maple						0.7				0.7	0.5	

Note: Data sets shown here were selected as being representative of trends in nine other data sets. Columns sum to 100. Bold values indicate overlap between groups with the same name.
*Class unique to each respective classification.

they are both weighted toward abundant and large canopy tree species (Radeloff et al. 1999). Relative importance may also be preferred in assessments of the value of regional habitat for certain wildlife species, which often respond to stem density in addition to tree diameters.

Ecological interpretation

Analysis of the statistical clusters and resulting maps reveals regional presettlement tree species associations and a complex patterning of forest types. Pine species, as previously noted by Finley (1951) and Curtis (1959), dominated on the three glacial sand plains within northern Wisconsin (Fig. 5, subsections 212Jm, 212Jk, and 212Ka). Composition within these regions, however, was by no means homogeneous. Jack pine was prevalent within the southern and central portions of the Bayfield Sand Plains (subsection 212Ka), the poorest soils in the study region, and substantial enclaves of jack pine – red pine, red pine, red pine – white pine, and white pine also existed (Fig. 5). The northern portion of this subsection was predominantly red pine. The other two sand plains, especially the Northern Highlands Pitted Outwash (subsection 212Jm), were dominated by red pine and white pine, contained a comparatively lower percentage of jack pine, and included a substantial proportion of aspen. The

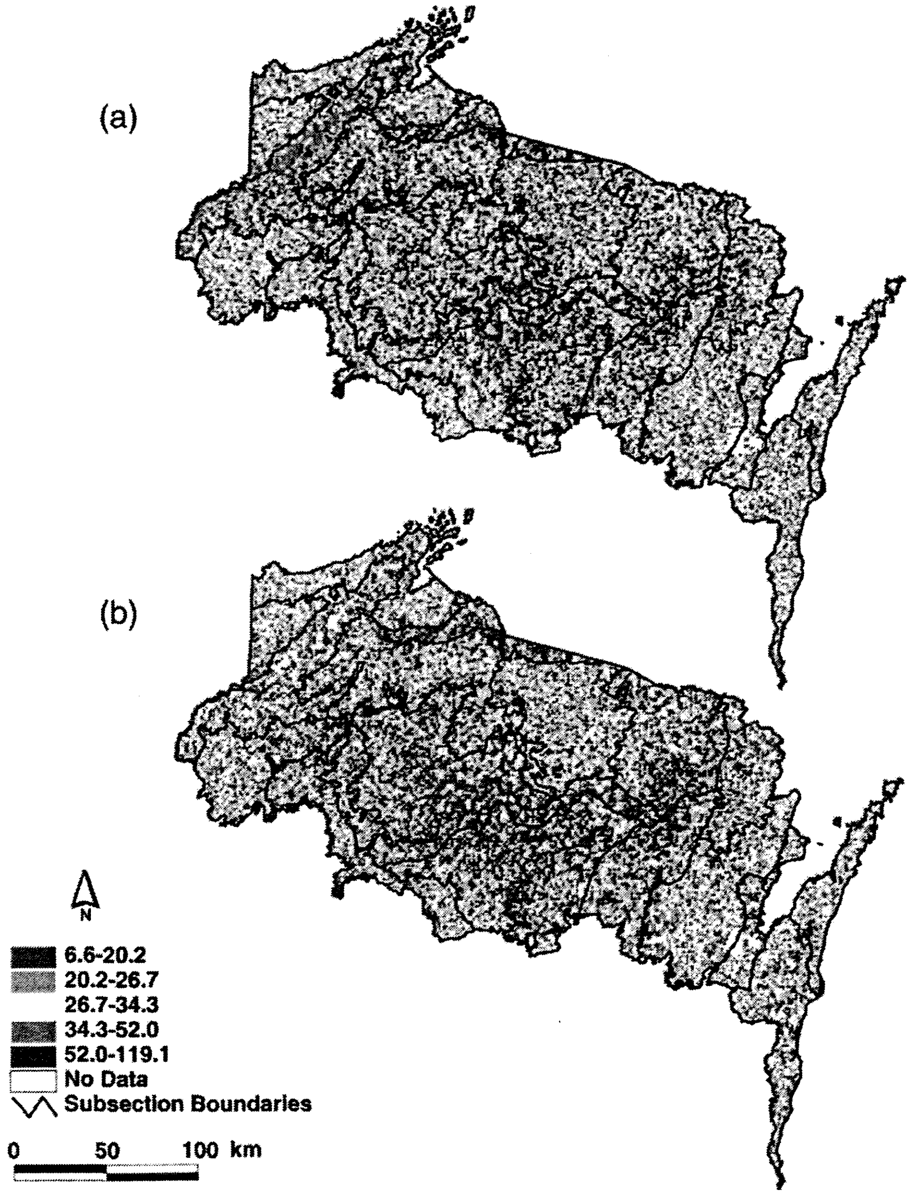
R. pine	R. pine – w. pine	Spruce – tamarack – w. pine – w. birch	S. maple	S. maple – w. pine*	S. maple – elm – b. ash – linden*	Tamarack	Tamarack – hemlock – y. birch – s. maple*	W. birch	W. oak	W. pine	Y. birch	Y. birch – s. maple
		1.7		1.1	2.0	1.7	0.9		1.3	0.6		
			1.8	2.2	2.0		0.9		7.6	1.6	0.7	0.7
		7.5	0.9					1.3		0.3		
							2.8					
			4.5	3.3			6.5					2.7
		4.2					3.7	1.3		1.3		
							7.4	1.3			3.7	1.3
0.9										0.3		
		4.2	7.1	20.7	72.0		1.9	6.5	1.3	2.2		8.0
				1.1	22.7		0.9		1.3	1.6		
83.0	1.7											
16.0	**95.8**	5.8				1.7		2.6	1.3	0.6		
		38.3				3.4	2.8			1.3		
		0.8	**83.0**	37.0	0.7			1.3				1.3
	1.7	19.2		1.1	0.7	**91.4**	65.7	1.3		1.3		
	0.8	4.2		1.1			0.9	**74.0**		1.3		
				1.1		1.7			**87.3**	1.9		0.7
		9.2		19.6						**84.8**		
							4.6	1.3		0.3	**93.3**	4.7
		5.0	2.7	12.0			0.9	9.1		0.6	2.2	**80.7**

classification based on relative importance also shows large patches of white birch forest within the Northern Highlands Pitted Outwash (subsection 212Jm).

More mesic conditions occurred on the moraines and till plains located in the central portion of the study area (subsections 212Jc, 212Je, 212Jf, 212Ji, 212Jj, 212Jl, and 212Js). These regions were covered by an intricate patterning of hemlock – northern hardwoods forest types. Although the traditional hemlock – northern hardwoods species were predominant (i.e., eastern hemlock, yellow birch, sugar maple, and linden), our analysis shows substantial variation within this generalized cover type. For example, yellow birch was especially predominant on the till plains (subsections 212Je and 212Js), where it is frequently mapped as a single-species class, and the presence of sugar maple was much more pronounced on the eastern well-drained moraines (subsections 212Jj and 212Jl). Interspersed within this hemlock – northern hardwoods matrix were sizeable patches of the white pine class, forested wetland (white-cedar, tamarack, and spruce – tamarack – white pine – white birch classes), as well as some aspen.

The glacial till and lake plains near Lake Michigan and Green Bay (subsections 212Hb, 212Hc, 212Hd, and 212He) were dominated by beech classes (beech, beech – sugar ma-

Fig. 6. Euclidean distance (unitless) from data point to cluster mean: (*a*) relative dominance and (*b*) relative importance. Classes were derived using natural breaks within the data.



ple, hemlock–beech), white oak, and white-cedar. Whereas the classification on dominance shows a homogenous beech class, the relative importance classification describes more variation in this cover class. Beech – sugar maple was more prevalent on the more southern Manitowoc Till Plain (212Hd) and hemlock–beech was more prevalent on the West Green Bay Till Plain (212Hb) and Door and Escanaba peninsulas and Lake Plain (212He). The dominance classification also shows a substantial proportion of northern pin oak – bur oak – black ash – sugar maple, which is not a very robust class (Table 4A). Looking at the individual cell values within Subsections 212Hb and 212He shows that black ash and sugar maple are greater contributors to this class than northern pin oak or bur oak. This is also suggested by the relative importance classification (Fig. 5*b*).

The St. Croix Moraine (212Jd) and the Lincoln Formation Till Plain (212Jg), located in the southwestern portion of the study area, are patterned with white pine, white oak, aspen, northern hardwoods (yellow birch, yellow birch – sugar maple, sugar maple – white pine, sugar maple – elm – black ash – linden), and northern pin oak – bur oak – black ash – sugar maple. Finley (1976), using subjective differentiation of generic bearing trees, designates much of the pine in this region as jack pine. Using our quantitative methodology, however, we have designated it as white pine, which coincides with the more mesic moraine and till plain substrates and the proximate northern hardwoods forests. The southern portion of these subsections lies within the "tension zone", a rough climatic boundary that is manifested as a plant species limiting line (Curtis 1959). The prevalence of oak classes (red oak, white oak, and bur oak – red oak – sugar maple – aspen) and aspen here likely attests to a higher frequency of fires. Fire disturbance was much more pronounced within the prairie–oak ecosystems of southern Wisconsin (Curtis 1959).

Despite distinct substrates, we have mapped very similar forest composition within the Lake Superior Clay Plain (subsection 212Ja) and the Mille Lacs Uplands (subsection 212Kb; Table 2). The lowlands within this region, which are extensive, were dominated by tamarack and spruce. Substantial proportions of white pine, aspen, and white birch classes were found in the uplands. Finley (1976) designates much of the Lake Superior Clay Plain as boreal forest, citing white spruce, balsam fir, tamarack, white-cedar, white birch, and aspen as prevalent species; however, these areas are statistically indistinguishable from the spruce – tamarack – white pine – white birch lowlands that existed in the pitted outwash plain and between the moraines in nearby subsections 212Jc, 212Jl, 212Jm, and 212Kb.

In summary, through the quantitative classification of regional forest types, we have revealed new associations within the PLS bearing-tree data. Because no a priori assumptions were made when deriving the classification, it is possible to analyze historical forest types against their environment and maps can be used to fulfill further scientific and management objectives.

Acknowledgements

We thank the Forest Landscape Ecology Laboratory group, especially Jason McKeefry, Volker Radeloff, and Ted Sickley, for helpful discussions and technical assistance provided during this work. Comments on earlier drafts of this manuscript were provided by Janine Bolliger, Emilie Grossman, Nancy Langston, Craig Lorimer, A-Xing Zhu, and two anonymous reviewers. The Wisconsin Department of Natural Resources has provided continuing funding for this work under the Pitmann-Robertson Program. Additional funding was provided by the U.S. Forest Service's North Central Research Station.

References

Albert, D.A. 1995. Regional landscape ecosystems of Michigan, Minnesota, and Wisconsin: a working map and classification. USDA For. Serv. Gen. Tech. Rep. NC-178.

Barrett, L.R., Leibens, J., Brown, D.G., Schaetzl, R.J., Zuwerink, P., Cate, T.W., and Nolan, D.S. 1995. Relationships between soils and presettlement forests in Baraga County, Michigan. Am. Midl. Nat. **134**: 264–285.

Batek, M.J., Rebertus, A.J., Schroeder, W.A., Haithcoat, T.L., Compas, E., and Guyette, R.P. 1999. Reconstruction of early nineteenth-century vegetation and fire regimes in the Missouri Ozarks. J. Biogeogr. **26**: 397–412.

Bourdo, E.A., Jr. 1956. A review of the General Land Office survey and of its use in quantitative studies of former forests. Ecology, **37**: 754–768.

Britton, N.L., and Brown, A. 1913. An illustrated flora of the northern United States, Canada and the British possessions, from Newfoundland to the parallel of the southern boundary of Virginia, and from the Atlantic Ocean westward to the 102d meridian. Vols. 1–3. 2nd ed. Charles Scribner's & Sons, New York.

Brown, D.G. 1998. Mapping historical forest types in Baraga County Michigan, USA, as fuzzy sets. Plant Ecol. **134**: 97–111.

Brown, R.T., and Curtis, J.T. 1952. The upland conifer–hardwood forests of northern Wisconsin. Ecol. Monogr. **22**: 217–234.

Christensen, E.M., Clausen, J.J., and Curtis, J.T. 1959. Phytosociology of the lowland forests of northern Wisconsin. Am. Midl. Nat. **62**: 232–247.

Cissel, J.H., Swanson, F.J., McKee, W.A., and Burditt, A.L. 1994. Using the past to plan the future in the Pacific Northwest. J. For. **92**: 30, 31, and 46.

Clark, J.S., and Royall, P.D. 1996. Local and regional sediment charcoal evidence for fire regimes in presettlement north-eastern North America. J. Ecol. **84**: 365–382.

Cleland, D.T., Avers, R.E., McNab, W.H., Jensen, M.E., Bailey, R.G., King, T., and Russell, W.E. 1997. National hierarchical framework of ecological units. *In* Ecosystem management: applications for sustainable forest and wildlife resources. *Edited by* M.S. Boyce and A. Haney. Yale University Press, New Haven, Conn. pp. 181–200.

Comer, P.J., Albert, D.A., Wells, H.A., Hart, B.L., Raab, J.B., Price, D.L., Kashian, D.M., Corner, R.A., and Schuen, D.W. 1995. Michigan's native landscape, as interpreted from the General Land Office surveys 1816–1856. Report to the U.S. Environmental Protection Agency, Water Division, Michigan Department of Natural Resources. Michigan Natural Features Inventory, Lansing, Mich.

Cottam, G., and Curtis, J.T. 1956. The use of distance measures in phytosociological sampling. Ecology, **37**: 451–460.

Cronon, W. 1983. Changes in the land: Indians, colonists, and the ecology of New England. Hill & Wang, New York.

Curtis, J.T. 1959. The vegetation of Wisconsin: an ordination of plant communities. University of Wisconsin Press, Madison, Wis.

Curtis, J.T., and McIntosh, R.P. 1951. An upland forest continuum in the prairie–forest border region of Wisconsin. Ecology, **32**: 476–496.

Davis, M.B. 1981. Quaternary history and the stability of forest communities. *In* Forest succession. *Edited by* D.C. West, H.H. Shugart, and D.B. Botkin. Springer-Verlag, New York. pp. 132–153.

Davis, M.B. 1984. Invasions of forest communities during the Holocene: beech and hemlock in the Great Lakes region. *In* Colonization, succession, and stability. *Edited by* A.J. Gray, M.J. Crawley, and P.J. Edwards. Blackwell Scientific Publications, London. pp. 373–393.

Davis, M.B., Woods, K.D., Webb, S.L., and Futyma, R.P. 1986. Dispersal versus climate: expansion of *Fagus* and *Tsuga* into the Upper Great Lakes region. Vegetatio, **67**: 93–103.

Davis, M.B., Sugita, S., Calcote, R.R., Ferrari, J.B., and Frelich, L.E. 1994. Historical development of alternate communities in a hemlock–hardwood forest in northern Michigan, USA. *In* Large-scale ecology and conservation biology. *Edited by* P.J. Edwards, R. May, and N.R. Webb. Blackwell Scientific Publishing, Boston, Mass. pp. 19–39.

Delcourt, H.R., and Delcourt, P.A. 1996. Presettlement landscape heterogeneity: evaluating grain of resolution using General Land Office Survey data. Landsc. Ecol. **11**: 363–381.

Digby, P.G.N., and Kempton, R.A. 1987. Population and community biology series: multivariate analysis of ecological communities. Chapman & Hall, London.

Everitt, B.S. 1980. Cluster analysis. 2nd ed. Heinemann, London.

Finley, R.W. 1951. The original vegetation cover of Wisconsin. Ph.D. dissertation, University of Wisconsin, Madison, Wis.

Finley, R.W. 1976. Original vegetation cover of Wisconsin. Map. USDA Forest Service, St. Paul, Minn.

Fries, R.F. 1951. Empire in pine: the story of lumbering in Wisconsin 1830–1900. Wisconsin State Historical Society, Madison, Wis.

Fulé, P.Z., Covington,W.W., and Moore, M.M. 1997. Determining reference conditions for ecosystem management of southwestern ponderosa pine forests. Ecol. Appl. **7**: 895–908.

Gartner, W.G. 1997. Four worlds without an Eden: pre-Columbian peoples and the Wisconsin landscape. *In* Wisconsin land and life. *Edited by* R.C. Ostergren and T.R. Vale. University of Wisconsin Press, Madison, Wis. pp. 331–350.

Gough, R.J. 1997. Farming the cutover: a social history of northern Wisconsin, 1900–1940. University Press of Kansas, Lawrence, Kan.

Grimm, E.C. 1984. Fire and other factors controlling the Big Woods vegetation of Minnesota in the mid-nineteenth century. Ecol. Monogr. **54**: 291–311.

He, H.S., Mladenoff, D.J., Sickley, T., and Guntenspergen, G. 2000. GIS interpolations of bearing tree records (1839–1866) for northern Wisconsin at multiple scales. J. Biogeogr. **27**: 1031–1042.

Karl, T.R., Williams, C.N., Jr., Quinlan, F.T., and Boden, T.A. 1990. United States Historical Climatology Network (HCN) serial temperature and precipitation data. Carbon Dioxide Information and Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn. Environ. Sci. Div. Publ. 3404.

Keys, J., Jr., Carpenter, C., Hooks, S., Koenig, F., McNab, W.H., Russell, W., and Smith, M.L. 1995. Ecological units of the eastern United States — first approximation. Map. USDA Forest Service, Washington, D.C.

Landres, P.B., Morgan, P., and Swanson, F.J. 1999. Overview of the use of natural variability concepts in managing ecological systems. Ecol. Appl. **9**: 1179–1188.

Lowell, K., and Jaton, A. (*Editors*). 1999. Spatial accuracy assessment: land information uncertainty in natural resources. Ann Arbor Press, Chelsea, Mich.

Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., McGwire, K.C., and Tinney, L.R. 1991. Remote sensing and geographic information system data integration: error sources and research issues. Photogramm. Eng. Remote Sens. **57**: 677–687.

Manies, K.L., and Mladenoff, D.J. 2000. Testing methods to produce landscape-scale presettlement vegetation maps from the U.S. public land survey records. Landsc. Ecol. **15**: 741–754.

Manies, K.L., Mladenoff, D.J., and Nordheim, E.V. 2001. Surveyor bias in forest data of the U.S. General Land Office records for northern Wisconsin. Can. J. For. Res. **31**: 1719–1730.

Manly, B. 1994. Multivariate statistical methods: a primer. 2nd ed. Chapman & Hall, London.

Marschner, F.J. 1974. The original vegetation of Minnesota. Map. USDA Forest Service, St. Paul, Minn.

McGarigal, K., Cushman, S., and Stafford, S. 2000. Multivariate statistical for wildlife and ecology research. Springer-Verlag, New York.

Milligan, G.W., and Cooper, M.C. 1987. Methodology review: clustering methods. Appl. Psychol. Meas. **11**: 329–354.

Mladenoff, D.J., and Pastor, J. 1993. Sustainable forest ecosystems in the northern hardwood and conifer forest region: concepts and management. *In* Defining sustainable forestry. *Edited by* G.H. Aplet, N. Johnson, J.T. Olson, and V.A. Sample. Island Press, Washington, D.C. pp. 145–180.

Mladenoff, D.J., Dahir, S.E., Nordheim, E.V., Schulte, L.A., and Guntenspergen, G.G. 2002. Narrowing historical uncertainty: probabilistic classification of ambiguously identified tree species in historical forest survey data. Ecosystems. In press.

Radeloff, V.C., Mladenoff, D.J., He, H.S., and Boyce, M.S. 1999. Forest landscape change: the northwest Wisconsin Pine Barrens before European settlement and today. Can. J. For. Res. **29**: 1649–1659.

SAS Institute Inc. 1990. SAS/STAT user's guide. Vol. 1. 4th ed. SAS Institute Inc., Cary, N.C.

Schulte, L.A., and Barnes, W.J. 1996. Presettlement vegetation of the lower Chippewa River Valley. Mich. Bot. **35**: 29–37.

Schulte, L.A., and Mladenoff, D.J. 2001. The original U.S. Public Land Survey Records: their use and limitations in reconstructing presettlement vegetation. J. For. **99**: 5–10.

Spencer, J.S., Jr., and Thorne, H.W. 1972. Wisconsin's 1968 timber resource — a perspective. USDA Forest Service, North Central Forest Experiment Station, St. Paul, Minn.

Stewart, L.O. 1935. Public land surveys — history, instructions, methods. Collegiate Press, Inc., Ames, Iowa.

Stone, R.N., and Thorne, H.W. 1961. Wisconsin's forest resources. USDA Forest Service, Lake States Forest Experiment Station, St. Paul, Minn.

Swain, A.M. 1978. Environmental changes during the past 2000 years in north-central Wisconsin: analysis of pollen, charcoal, and seeds from varved lake sediments. Quat. Res. **10**: 55–68.

Thomas, J.W. 1996. Forest Service perspective on ecosystem management. Ecol. Appl. **6**: 703–705.

Vale, T.R. 1998. The myth of the humanized landscape: an example from Yosemite National Park. Nat. Areas J. **18**: 231–236.

Wallin, D.O., Swanson, F.J., Marks, B., Cissel, J.H., and Kertis, J. 1996. Comparison of managed and pre-settlement landscape dynamics in forest of the Pacific Northwest, USA. For. Ecol. Manage. **85**: 291–309.

Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. **58**: 236–244.

Webb, T., III. 1974. A vegetational history from northern Wisconsin: evidence from modern and fossil pollen. Am. Midl. Nat. **92**: 12–33.

White, M.A., and Mladenoff, D.J. 1994. Old-growth forest landscape transitions from pre-European settlement to present. Landsc. Ecol. **9**: 191–205.

Wisconsin Department of Natural Resources (WDNR). 1999. Ecological units of Wisconsin — first approximation, including LandType Associations (LTAs). Wisconsin Department of Natural Resources, Madison, Wis.

Zhang, Q.F., Pregitzer, K.S., and Reed, D.D. 2000. Historical changes in the forests of the Luce District of the Upper Peninsula of Michigan. Am. Midl. Nat. **143**: 94–110.