Journal of Biogeography, **27**, 1031–1042

GIS interpolations of witness tree records (1839–1866) for northern Wisconsin at multiple scales

Hong S. He, David J. Mladenoff, Theodore A. Sickley and Glenn G. Guntenspergen[1]
Department of Forest Ecology and Management, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53706, and [1]U.S. Geological Survey, Biological Resources Division, Patuxent Wildlife Research Center, 11510 American Holly Dr., Laurel, MD 20708 U.S.A.

Abstract

To construct forest landscape of pre-European settlement periods, we developed a GIS interpolation approach to convert witness tree records of the U.S. General Land Office (GLO) survey from point to polygon data, which better described continuously distributed vegetation. The witness tree records (1839–1866) were processed for a 3-million ha landscape in northern Wisconsin, U.S.A. at different scales. We provided implications of processing results at each scale. Compared with traditional GLO mapping that has fixed mapping scales and generalized classifications, our approach allows presettlement forest landscapes to be analysed at the individual species level and reconstructed under various classifications. We calculated vegetation indices including relative density, dominance, and importance value for each species, and quantitatively described the possible outcomes when GLO records are analysed at three different scales (resolution). The 1×1-section resolution preserved spatial information but derived the most conservative estimates of species distributions measured in percentage area, which increased at coarser resolutions. Such increases under the 2×2-section resolution were in the order of three to four times for the least common species, two to three times for the medium to most common species, and one to two times for the most common or highly contagious species. We mapped the distributions of hemlock and sugar maple from the pre-European settlement period based on their witness tree locations and reconstructed presettlement forest landscapes based on species importance values derived for all species. The results provide a unique basis to further study land cover changes occurring after European settlement.

Keywords
Presettlement, witness tree, GLO record, GIS interpolation, spatial resolution, forest landscape, Wisconsin.

INTRODUCTION

Records from the U.S. General Land Office (GLO) surveys conducted from the late 1700s to early 1900s provide a key data source of forest and environmental conditions of the pre-European settlement period (Bourdo, 1956). Under the GLO, the public land survey system (PLSS) was developed, in which land was divided into a grid of square townships, each containing 36 square mile (1.6×1.6 km) sections (Fig. 1). At each section corner and the mid-point between section corners (quarter corner) a survey marker, usually a stake, was placed. Between two and four witness trees were identified and measured to document the corner location. Tree species, diameter, distance from the corner, and compass bearing were recorded in the surveyors' notebook. When traversing the section lines, the surveyors also recorded trees that fell along the section lines (line trees), and trees around survey posts set where section lines intersected lakes or impenetrable

Correspondence: Hong S. He, School of Natural Resources, University of Missouri-Columbia, 203 ABNR Building, Columbia, MO 65211. E-mail: HeH@missouri.edu

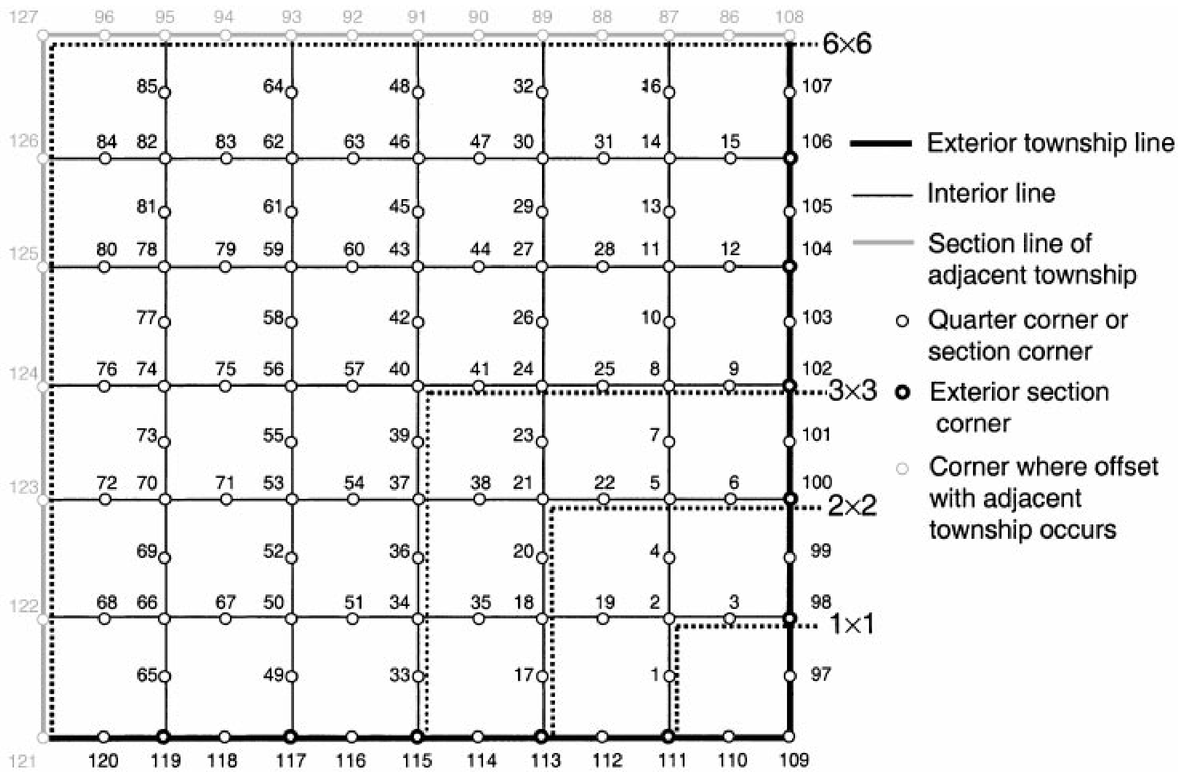

Figure 1 Interior and exterior point identification numbers for a township and the point sampling scheme under 1 × 1, 2 × 2, 3 × 3, and 6 × 6-section resolution, respectively.

wetlands (meander corners). These field notebooks were later microfilmed for long-term storage and public access. For the state of Wisconsin, the total number of individual trees recorded is estimated to be near 500,000. These large data sets and the regular and systematic locations of the survey points make the GLO survey records very important for reconstructing and analysing forests and environments of the pre-European settlement period.

Forest density, species composition and dominance are the most important information that can be derived from GLO survey records (e.g. Galatowitsch, 1990; Fralish *et al.*, 1991; Leitner *et al.*, 1991; Nelson, Redmond & Sparks, 1994; Nelson, 1997). For each GLO survey point or a group of points falling in a study area, forest tree density is estimated from recorded distances to witness trees using established distance-based sampling methods (Cottam & Curtis, 1956); tree species relative density (% species composition) is estimated from species counts; and dominance (% basal area) is calculated from recorded species diameters. Since each GLO record corresponds to a unique spatial location (section corner or quarter corner), a GLO data set can be easily imported into a geographical information system (GIS) as a point coverage. These data points are often directly plotted to describe the presettlement forest of an area. However, a point coverage is not sufficient to describe vegetation that is continuously distributed over the landscape, and more relevant data forms such as grids or polygon coverages (maps), interpolated from the survey points, are desirable (He & Ventura, 1995; Brown, 1998a, b; White & Mladenoff, 1994).

Prior to the development of GIS and spatial analysis techniques, interpolating GLO survey records from individual points to polygons of various vegetation types was a manual mapping processes, and has proven difficult; (Gordon, 1969; Anderson, 1970; Finley, 1976). Using GLO survey records often involves the manual and time-consuming process of transcribing the surveyors' notebooks. Individual studies have typically been limited to small areas comprising 50–500 GLO survey points (e.g, Lorimer, 1977; Kapp, 1978; Mladenoff & Howell, 1980; Whitney, 1982, 1990; Iverson, 1988; Abrams & Ruffner, 1995; Delcourt & Delcourt, 1996). To map large areas, such as the state of Wisconsin, it was infeasible to manually compute vegetation indices based on all GLO survey points and then delineate vegetation boundaries. The mapping method commonly used was to subjectively delineate vegetation patches based on GLO tree distributions, characteristics, and surveyor descriptions. Numerous subjective decisions were made where vegetation boundaries were indistinct. Where many species combinations occurred, generalized vegetation representations were used. Although these time-consuming manual methods have created useful vegetation maps in the past, there are significant limits associated with these methods: (1) the fixed, single mapping scale (e.g. 1 : 500,000 map of original vegetation of Wisconsin, Finley, 1976) may exclude the examination of phenomena occurring at other scales;

(2) the fixed and often general vegetation classification makes it difficult to address issues at the species level; and (3) the maps are difficult or impossible to recreate or compare with other areas due to the subjectivity inherent in the manual mapping process.

To reconstruct the presettlement forest landscape, we have developed a GIS processing approach that overcomes the limits of traditional mapping. We use a moving sampling window (polygon) to interpolate vegetation data of GLO survey plots to polygons, with the smallest window size (1 × 1 section) in accordance with the GLO survey resolution. This processing approach preserves all original information (including survey bias) in the GLO data set. In addition, different processing window sizes such as 2 × 2, 3 × 3, or 6 × 6 sections can be chosen so that GLO records are analysed at a resolution of interest. Different processing resolutions may lead to different results since vegetation indices such as species presence/absence and dominance (basal area/unit area) derived from GLO survey records are resolution dependent. The changes in these indices among resolutions are often found to be non-linear (e.g. Kunin, 1998), and often affected by species' spatial configurations on a landscape. Therefore, understanding the thresholds of vegetation indices under various resolutions is important when interpreting results of GLO data for a given resolution. Thus, in this paper, we endeavour to construct a presettlement forest landscape in a large region in northern Wisconsin, U.S.A., and examine species abundance and dominance with a respect of spatial resolutions.

METHODS

GIS processing

GLO records were transcribed into a GIS-compatible database while an error checking and quality control procedure was applied (Manies, 1997). They were then transformed into an Arc/Info point coverage that contains all fields recorded in the GLO survey. In our GIS processing approach, we calculated vegetation indices including relative and absolute density, relative dominance and relative importance, at four moving window sizes: 1 × 1, 2 × 2, 3 × 3 and 6 × 6 sections (Fig. 1). The 1 × 1 window covers one section, 2 × 2 four sections, 3 × 3 nine sections, and 6 × 6 one township. Window sizes 4 × 4, 5 × 5 and those larger than 6 × 6 have not been implemented, because it would require simultaneous processing of multiple townships, substantially increasing the programming tasks without significant gain in the ecological context. The current moving window program (MWINDOWS) utilizes regular points belonging to a given window (points under grey, dashed lines in Fig. 1) to calculate vegetation indices for that window. For example, for a 6 × 6 window, or a township window, all 108 regular survey points are incorporated as processing points within that township (Fig. 1). The same rule applies to all subsequently smaller window sizes and this ensures that witness trees at each point are used only once, even though the section boundaries are shared (Fig. 1).

MWINDOWS is comprised of a set of Arc/Info AML programs (ESRI Inc., 1996) and is able to perform preprocessing vegetation reclassification by allowing users to define appropriate vegetation classes of interest. For example, users can specify one or more species of interest to process and lump the remaining species into a dummy class. Users can also classify species into classes that are at single species level, genus group level, or more general levels such as hardwood, conifer, upland and lowland. This reclassification feature is useful since some species are identified at both species and genus levels or only genus in the survey notes, such as paper birch, yellow birch and birch (Manies, 1997; Nelson, 1997; Mladenoff *et al.*, 1999).

Witness tree data processing

For each window at any size, the program determines presence/absence for each species or a user-defined class (in reclassification file). If a species is present, the area of the processing window is included in the total area in which the species occurs. Thus, the percentage area or proportion of the landscape occupied by a given species or class i, $Area\%\_C_i$, is defined as:

$$Area\%\_C_i = areaSum\_C_i / totalArea \cdot 100$$

where $areaSum\_C_i$ is the total area containing species or class i, and $totalArea$ is the total area of the landscape. Therefore, $areaSum\_C_i$ under the 1 × 1-section resolution can maximally increase to $4 \times areaSum\_C_i$ under the 2 × 2-section resolution (four times); $areaSum\_C_i$ under the 2 × 2-section resolution can maximally increase to $9/4 \times areaSum\_C_i$ under the 3 × 3-section resolution (2.25 times).

For each window at any size, the program also calculates absolute density, relative density and relative basal area for each species or class to derive relative importance value. Absolute density is calculated as the average densities measured for the first witness trees, based on a method developed by Cottam & Curtis (1956). Incorporating second or more witness trees for such a calculation has been found more biased (Leitner *et al.*, 1991). Relative density of species or class i, $Density\_C_i$, is calculated as:

$$Density\_C_i = \sum_{j=1}^{m} C_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{m} C_{ij} \cdot 100\%,$$

where n is the number of species or classes in the processing window, m is the number of individuals of a given species or class in the window, and C_{ij} is individual j of species or class i. $Density\_C_i$ ranges from 0 to 100%.

Relative basal area of a species or a class i, Ba_C_i, is calculated as:

$$Ba\_C_i = \sum_{j=1}^{m} BasalArea_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{m} BasalArea_{ij} \cdot 100\%,$$

where n and m are the same as described above. $BasalArea_{ij}$ is basal area of individual j of species or class i. Ba_C_i ranges between 0 and 100%.

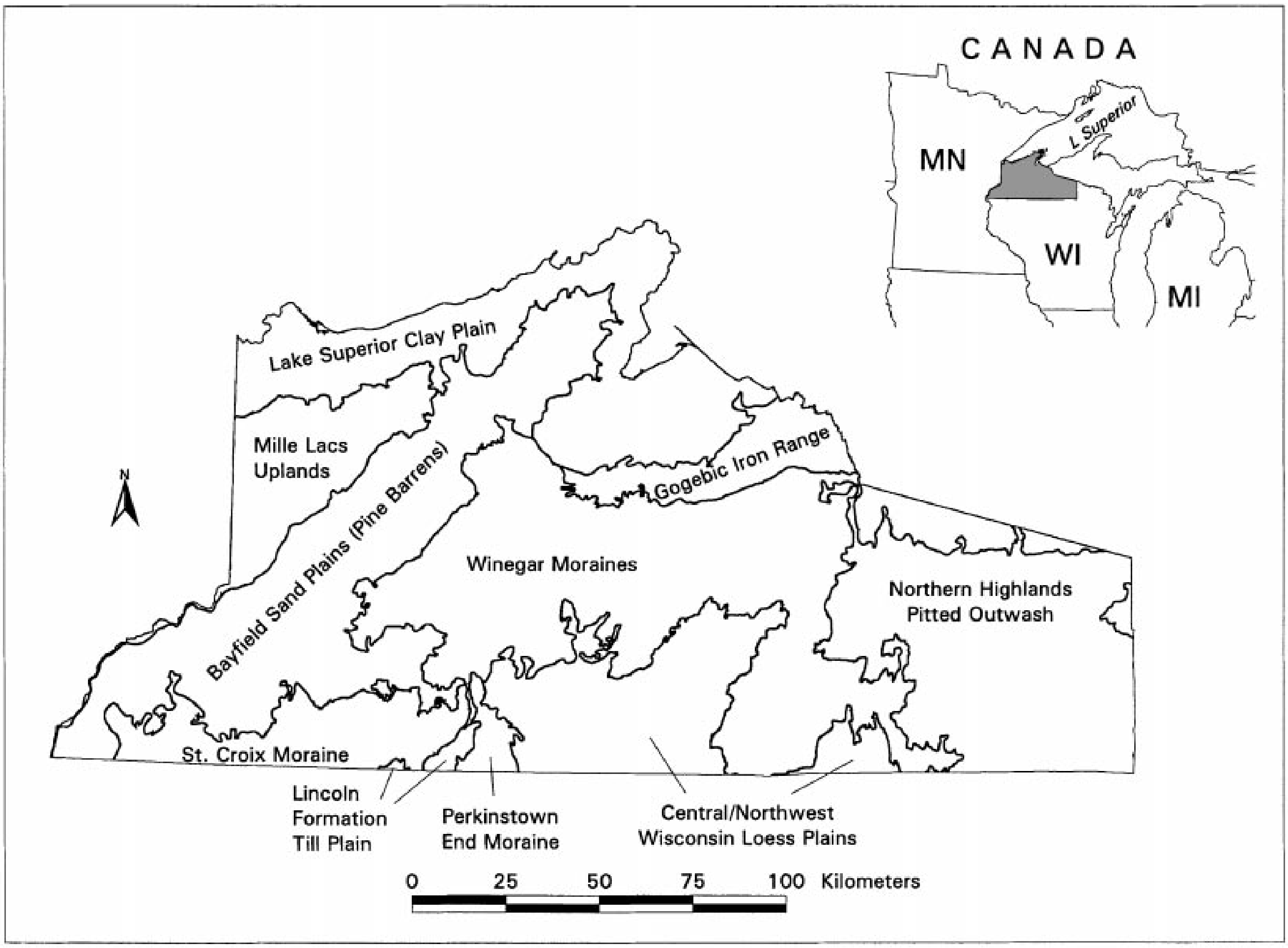

Figure 2 Outline of the study area containing about 3 million ha in northern Wisconsin, U.S.A., and the major ecological units (ecoregions) in the study area.

Relative importance value of a species, RIV_C_i, is calculated as the average of relative density and relative basal area as:

$$RIV\_C_i = (Density\_C_i + Ba\_C_i)/2$$

in which RIV_C_i ranges from 0 to 100%, with small numbers indicating low levels of importance. *RIV* indicates the relative level of importance of a species in relation to other species that occur in the same window. For the entire landscape, the number of windows with a species *RIV* at any given level can be queried. In other words, area where species occur can be plotted against *RIV* for each species. This allows us to assess species importance across the landscape. Result of querying $RIV > 0$ for a given species returns the total area of that species in the landscape at the given resolution.

RIV provides a basis for comparing species within each window. Therefore a variety of forest classification schemes can be developed using quantitative criteria. In our effort to reconstruct the presettlement forest landscape, we use the species with the highest *RIV* to represent each window.

STUDY AREA

Our study area covers over 3 million hectares of northern Wisconsin, from township 36–53 north and ranges 11 east to 20 west (Fig. 2). GLO survey for this area was conducted between 1839 and 1866. The area contains about 300 townships, 12,258 sections, about 36,000 corners (section and quarter), and more than 81,000 individual tree records. The study area is located in a transitional zone between deciduous forest to the south and mixed deciduous and evergreen forest to the north (Curtis, 1959). Quaternary geological events and climatic variations shaped the physiography and soils of this large area (Keys *et al.*, 1995; Host *et al.*, 1996), which vary from nutrient-poor sandy soils or outwash, such as in the sand barrens in the west; clay and silty soils along Lake Superior; and loam soils on moraines and till plains (Fig. 2). Historical information on vegetation of this region is sparse. Little is know about the land use of indigenous Americans. The original forested area was heavily logged around 1900. The forests now are almost all secondary and third-growth and aged between 40 and 90 years old (He *et al.*, 1998). Currently, the dominant species include sugar maple (*Acer saccharum*), quaking aspen (*Populus tremuloides*), big-toothed aspen (*P. grandidentata*), paper birch (*Betula papyrifera*), yellow birch (*B. alleghaniensis*), northern red oak (*Quercus ruba*), red pine (*Pinus resinosa*), jack pine (*P. banksiana*), and balsam fir (*Abies balsamea*) (Mladenoff & Pastor, 1993).

RESULTS

For our study area, there are approximately 12,000 sections at the 1 × 1-section resolution, 3000 at 2 × 2, 1000 at 3 × 3, and 300 at 6 × 6. Since the spatial resolution of the 6 × 6 section is relatively coarse (6 × 6 mile, 9.6 × 9.6 km) for our study area, we only present results from the three finer resolutions. We analysed eight common upland tree species: white pine (*P. strobus*), eastern hemlock (*Tsuga canadensis*), red pine, jack pine, sugar maple, aspen, balsam fir, white ash (*Fraxinus americana*), and northern red oak. Other species were included in calculation of the vegetation indices, but they are not discussed since they are either not significant (small proportions, <1%) or not well identified to the species level in the GLO survey, such as paper birch and yellow birch.

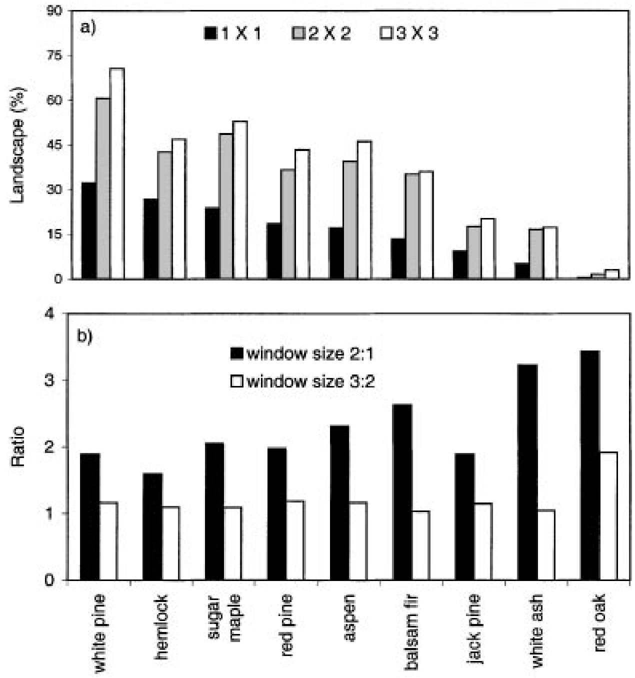

Figure 3 (a) The percentage area of eight upland tree species under three resolutions for northern Wisconsin. (b) The ratios of percentage changes of the section resolutions: 1 × 1 to 2 × 2 and 2 × 2 to 3 × 3.

Presettlement species percentage area and distribution

Species percentage area, the proportion of the landscape on which a species occurred, was derived for all eight selected species under three resolutions. At the 1 × 1-section resolution, the species with the largest percentage area in the presettlement period was white pine occurring on 32% of the landscape, followed by hemlock, 27% and sugar maple found on 24% of the landscape (Fig. 3a). Red pine, aspen and balsam fir were mid-abundant species found on 13–19% of the landscape, while jack pine was found on about 9% of the total landscape (Fig. 3a). White ash (*Fraxinus americana*) and northern red oak were not common, occurring on 5 and 0.5% of the landscape, respectively (Fig. 3a). Individual species of the presettlement period can be further analysed by mapping their spatial distributions. For example, hemlock was found occurring primarily on moraines and loess plains in the central part of the study region and reached its western range limit at the eastern boundaries of the sand barrens ecoregion (Fig. 4a). Sugar maple (Fig. 4b), however, was widely distributed on most ecoregions, except outwash and sand plains, and coexisted with hemlock in the presettlement distributions. At the 1 × 1-section resolution, both species appeared not to exist in large patches especially for sugar maple (Fig. 4b).

At the 2 × 2-section resolution, white pine distributed on 60% of the landscape. Sugar maple surpassed hemlock and became the species of second largest area, increasing to 49% of the landscape, while area covered by hemlock increased to 42% of the landscape. This change can also be seen in their spatial distributions as we observe that patch sizes of both species increase and their geographical distribution ranges become obvious (Fig. 4c–d). The order of area covered by the remaining species remained unchanged, but the magnitude of increase in percentage area under the 2 × 2-section resolution varied significantly among these species (Fig. 3a). As explained previously, the maximum possible increase in species percentage area from 1 × 1 to 2 × 2 section is four times (from one to four sections). The observed trend of such increases was that the least common species had the largest increases (Fig. 3b). This is contrary to what typically happens when aggregating raster data, where the area of uncommon classes, especially if broadly distributed, decrease with a decrease in spatial resolution (larger window sizes). The largest increases (three to four times) in percentage areas were found for the least common species, northern red oak and white ash, which increased to 3.4 and 3.2 times of their percentage areas from the 1 × 1–2 × 2-section resolution (Fig. 3b). Modest increases (two to three times) in percentage area were found for the mid-abundant species such as balsam fir, aspen and red pine, which increased their percentage areas by 2.6, 2.3 and 2.0 times from the 1 × 1–2 × 2-section resolution, respectively (Fig. 3b). The smallest increases (<2.0) were found for hemlock and jack pine, which increased by 1.5 and 1.8 times, respectively.

At the 3 × 3-section resolution, white pine, sugar maple and hemlock percentage areas increased to 71, 53, and 47%

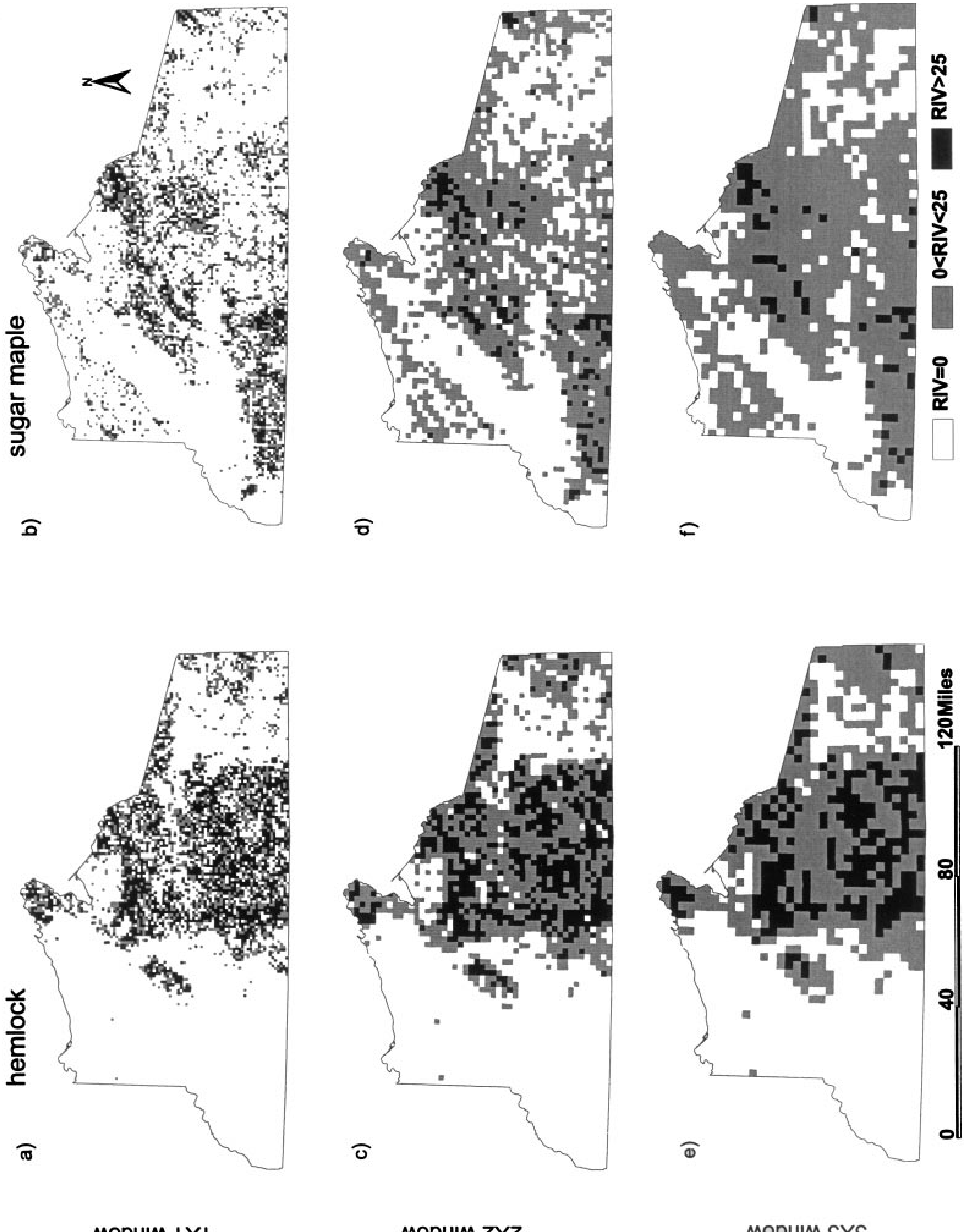

Figure 4 Presettlement hemlock and sugar maple distribution under the 1 × 1, 2 × 2 and 3 × 3-section resolutions in north Wisconsin. The lighter grey shading indicates the species present and *RIV* < 25; the darker grey shading indicates the species present and RIV ≥ 25.

of the landscape, remaining as the species of largest percentage areas (Fig. 3b). The percentage area for all the remaining species increased but the magnitudes of such increases were much less than those observed for the 1×1–2×2-section resolution (Fig. 3b). For example, the levels of spatial aggregation were more similar between the 2×2 and 3×3-section resolution than between the 1×1 and 2×2-section resolution for both hemlock (Fig. 4e) and sugar maple (Fig. 4f). The maximum possible increase in species abundance from 2×2 to 3×3 section was 2.25 times (from four to nine sections). However, with the exception of the uncommon, but widely distributed, northern red oak with a 1.9-times increase, the increases for the remaining species were relatively small (< 1.2, where 1.0 indicates no increase), regardless of their percentage areas under the 2×2 section (Fig. 3b). This suggests that the thresholds of species percentage area have been reached under the 2×2-section resolution given actual species distributions and the sampling (survey corner) density. Measuring the levels of spatial aggregation using aggregation index (He, *et al.*, in press) has confirmed such results.

Presettlement species dominance and the forest landscape

For a given resolution, *RIV*s were derived for each species, based upon the number of occurrences, density and tree sizes recorded in the GLO data set. Summarizing *RIV*s for a given species at a given window size by aggregating individual windows to the entire landscape provides an overall assessment of the species' importance across the landscape at that resolution. This can be done by deriving *RIV* curves, the proportion of the landscape covered by the species against *RIV*s (e.g. Fig. 5a). The total proportion of the landscape over which a species occurs at a given *RIV* level decreases as *RIV* increases. Therefore, the areas covered by a species at low *RIV*s are always larger than the areas covered by that species at high *RIV*s. The greater an area over which a species maintains high *RIV*s, the greater the importance of that species. Prior to European settlement, white pine and hemlock had high *RIV*s compared with other species, indicating their overall importance in the region (Fig. 5). Sugar maple was very common but its *RIV* curve decreased rapidly as *RIV* increased (Fig. 5a). This indicates that sugar maple was less important than white pine and hemlock in the presettlement landscape. Jack pine showed strong local importance since it did not occur over a large proportion of the landscape at $RIV \geq 0$, but maintained relatively high proportions on the landscape over all *RIV*s (Fig. 4a). Ash and balsam fir showed the lowest importance, as their curves decreased most rapidly at increasing *RIV* levels (Fig. 4a).

We constructed the presettlement forest landscapes using the species with highest *RIV* to represent important forest types for each window. Important forest types and their proportions on the landscape were summarized for each spatial resolution (Table 1). White pine, hemlock, birch (yellow and paper birch), red pine, tamarack, jack pine, sugar maple and aspen forests were found to be the most common important types accounting for 81% of the region under the 1×1-section resolution, 91% under the 2×2-section resolution and 95% under the 3×3-section resolution. More than 25 other species and classes accounted for the rest of the percentage. They are mapped as a lumped class, with statistics of each class provided (Table 1).

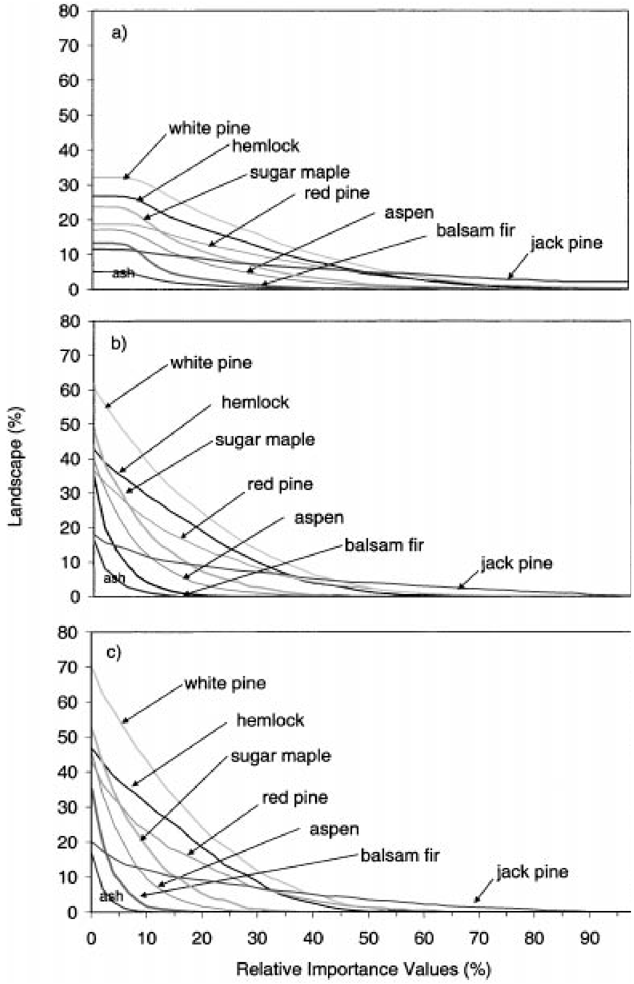

Figure 5 Species relative importance values (*RIV*s) vs. the percentage areas occupied under the (a) 1×1, (b) 2×2 and (c) 3×3-section resolutions ($0 < RIV \leq 100$).

Jack pine was important on almost 8% of total landscape occurring largely on the central and southern sand barrens in the west of the region (Fig. 6a). Red pine was important on the northern sand barrens as well as a pitted outwash ecoregion in the east, while white pine, the most important species in this region, occurred largely on better outwash soils than those of jack and red pines, and on moraine and clay plain ecoregions surrounding jack pine and red pine (Fig. 6a). Hemlock is the second most dominant species in this region. It occurred between the white pine-dominated forest to the west and red pine-dominated outwash to the east. Birch, not differentiated to the species level, but largely

Table 1 Presettlement dominant forest types and their percentage areas on the study region under three processing resolutions.

		1 × 1 section		2 × 2 section		3 × 3 section	
		No. of section	%	No. of section	%	No. of section	%
White pine	*Pinus strobus*	1989	17.11	2540	21.85	2995	25.77
Hemlock	*Tsuga canadensis*	1534	13.20	2110	18.15	2497	21.48
Birch	*Betula*	1512	13.01	1635	14.07	1616	13.90
Tamarack	*Larix laricina*	1212	10.43	1065	9.16	1017	8.75
Red pine	*Pinus resinosa*	1120	9.64	1279	11.00	1244	10.70
Jack pine	*Pinus banksiana*	806	6.93	875	7.53	923	7.94
Sugar maple	*Acer saccharum*	745	6.41	683	5.88	503	4.33
Aspen	*Populus*	531	4.57	351	3.02	210	1.81
Cedar	*Thuja occidentalis*	385	3.31	263	2.26	189	1.63
Spruce	*Picea glauca*	376	3.23	259	2.23	196	1.69
Pine	*Pinus*	158	1.36	134	1.15	107	0.92
Balsam fir	*Abies balsamea*	155	1.33	44	0.38	16	0.14
White spruce	*Picoa glauca*	153	1.32	83	0.71	50	0.43
Maple	*Acer*	146	1.26	52	0.45	36	0.31
White birch	*Betula papyritera*	144	1.24	89	0.77	45	0.39
White oak	*Quercus alba*	136	1.17	103	0.89	75	0.65
Elm	*Ulmus*	119	1.02	45	0.39	63	0.54
White cedar	*Thuja occidentalis*	74	0.64	45	0.39	23	0.20
Ash	*Eraxinus*	69	0.59	7	0.06		0.00
Basswood	*Tilia*	68	0.59	14	0.12		0.00
Black oak	*Q. velutina*	66	0.57	20	0.17	20	0.17
Red oak	*Q. ruba*	34	0.29	12	0.10		0.00
Bur oak	*Q. macrocarpa*	26	0.22	31	0.27	27	0.23
Willow	*Salix*	26	0.22	10	0.09	7	0.06
Pin oak	*Q. ellipsoidalis*	14	0.12	8	0.07		0.00
Ironwood	*Dialium*	7	0.06		0.00		0.00
Oak	*Quercus*	6	0.05		0.00		0.00
White ash	*Fraxinus americana*	5	0.04		0.00	1	0.01
Walnut	*Juglans*	2	0.02	4	0.03		0.00
No tree	–	2	0.02		0.00		0.00
Hickory	*Carya*	1	0.01		0.00		0.00
lowland	–	1	0.01		0.00		0.00

%, percentage landscape. Birch represents undifferentiated paper birch (*Betula papyrifera*) and yellow birch (*B. alleghaniensis*); aspen represents undifferentiated trembling aspen (*Populus tremuloides*) and big-toothed aspen (*P. grandidentata*); pine represents undifferentiated white pine, jack pine, and red pine.

composed of yellow birch in this region (Mladenoff *et al.*, 1999), was often associated with hemlock in the central part of the region. Sugar maple, was a widely distributed, common species, but tended to be of secondary importance at the presettlement period. It ranked third in percentage area of occurrence (24%) but was important only on 6% of the landscape under the 1 × 1-section resolution. It was also noticeable that tamarack (Larix laricina), a lowland species, was fairly common in this region (Fig. 6a).

Species *RIV* and the presettlement forest landscape can be examined under different resolutions (Fig. 5b–c). The area over which a species occurs at a given *RIV* also varies by window sizes. The *RIV* curves of the 1 × 1 section start lower than the two other-section resolutions since the proportion of the landscape covered by a species at low *RIV*s approaches their occurrence on the landscape, which are the lowest in percentage area among the three-section resolutions. But the 1 × 1 curves overtake the other two resolutions as *RIV* increases, as seen for hemlock (Fig. 5a) and sugar maple (Fig. 5b). This is because in the 1 × 1-section resolution, a maximum of six to eight trees was recorded, therefore within-window species heterogeneity is low and *RIV*s are generally high. Conversely, the 2 × 2 and 3 × 3-section resolutions capture more trees (up to 64 in a 3 × 3 window), and more species heterogeneity than the 1 × 1 section. Therefore, the corresponding *RIV* curves are generally low, suggesting that no single species become absolutely dominant. Relatively small differences were found for *RIV* curves of the 2 × 2 and the 3 × 3 sections. This suggests that thresholds of species heterogeneity are captured with the 2 × 2 section window size and increasing to the 3 × 3 section does not substantially incorporate more species (Fig. 4a–f).

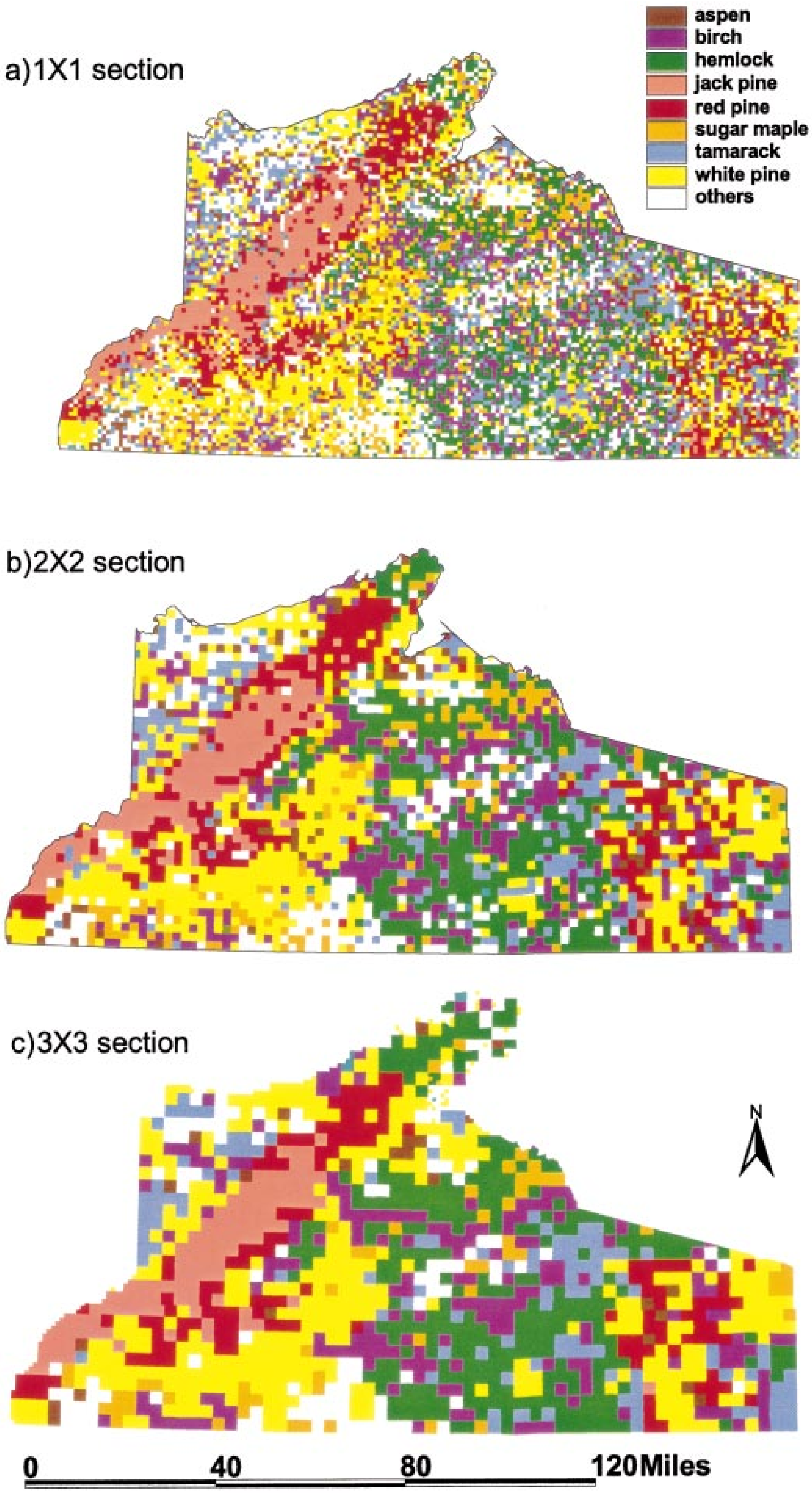

Figure 6 Constructed presettlement forest landscapes under (a) 1×1, (b) 2×2 and (c) 3×3-section resolutions for north Wisconsin.

DISCUSSION

GLO records of species occurrence and the processing resolutions determine the mapped species distribution. The locations of species recorded in the GLO data set form spatial patterns that directly affect the magnitudes of mapped abundance increases for a given species when moving to coarser spatial resolutions. Widely distributed species such as white pine and sugar maple have more surrounding area that can be added to the percentage area measurement when processed using coarser spatial resolutions. On the other hand, the percentage area of species that occur in highly contagious or aggregated patterns to begin with will not increase as dramatically under increasing window size. Hemlock, for example, occurred primarily in the central part of the study region and is largely found as large contiguous patches (Fig. 4a). This was further quantified using an aggregation index (He *et al.*, in press). Sugar maple (Fig. 4d), however, is widely distributed in this region and does not exist in patches as large as hemlock. Even though hemlock percentage area is higher than sugar maple under the 1×1-section resolution (Fig. 5a,d), we observed larger increases for sugar maple from the 1×1–2×2-section resolution than those for hemlock (Fig. 5b,e). Jack pine distribution is similar to hemlock and highly contiguous or locally dominant in the sand barrens of north-western Wisconsin and the sandy outwash soils in the east. Its increase of 1.8 times when moving from the 1×1 to the 2×2-section resolutions is second lowest after hemlock (Fig. 3b). Red pine occurs more frequently with a less contiguous distribution along the margins of the barrens. An increase of 2.0 times is observed for red pine under the 2×2-section resolution, larger than that of jack pine (Fig. 3b).

Unlike current forest survey data sets such as Forest Inventory and Analysis data (Hansen *et al.*, 1992), the purpose of the GLO survey was not forest inventory. The 1×1 section uses three survey corners totalling six to eight trees to represent the window. Thus, only a limited number of trees were recorded at each corner and some species likely occurred but were not recorded. Therefore, the 1×1-section resolution produces the most conservative estimates (likely underestimates) of all species presence distributions. Mapped species distributions as percentage area increase when moving to a 2×2-section resolution simply due to the larger area that the sample represents. This may compensate for the conservative estimate from the 1×1 section. Species percentage area reached certain thresholds at the 2×2 section, which do not increase substantially when moving to a coarser resolution, as we observed the smallest changes in species percentage area when moving from the 2×2–3×3 section (Fig. 3b). In fact, the processing results under the 2×2-section resolution approximate the species distributions reported in other studies in this region at the similar scales (Curtis, 1959; Finley, 1976). Furthermore, the similarity of the *RIV* curves of the 2×2 and 3×3 sections indicate that moving to window size larger than 2×2 sections sacrifices spatial resolution without capturing greater species heterogeneity. However, results from larger window sizes are suitable for studies at broad scales, since large window sizes decrease the patchiness of species distribution and make applications such as the delineation of species distribution at regional scales possible. Users may find that results of the 3×3-section resolution are desirable when comparing the GLO data with land-type association (LTA) map units (< 1 : 1,000,000, Keys *et al.*, 1995) since fragmented species distributions seen in the 1×1 and 2×2-section resolutions are aggregated into large patches suitable for the scales of regional planning or management purposes (e.g. Fig. 5c,f).

Various survey biases exist in the GLO survey data set, such as biases of selecting tree species (Buordo, 1956; Manies, 1997). Surveyors were instructed or tended to mark species that were medium in diameter, mature and long-lived. Therefore, small and/or uncommon species tend to be less represented than the common and dominant species. More biases may be inherited when mapping the presence or absence of minor species such as ash, oak, fir and aspen than mapping common species. On the other hand, because of the biases towards dominant species, construction of presettlement forest landscapes from dominant forest types (e.g. Fig. 6) produces more reliable maps than maps of species presence or absence (e.g. Fig. 4). Nevertheless, mapping of species presence or absence with the GLO data provides a unique reference of historical distributions that otherwise does not exist.

Compared with dominant forest types of the presettlement period (Table 1), the forest landscape in northern Wisconsin today is very different in both species distribution and dominant forest types. After heavy logging, first for conifers, then for all other species, the white pine- and hemlock-dominated presettlement landscape has been replaced with deciduous and mixed forests (Curtis, 1959). Sugar maple, a historically abundant and late successional species, also experienced extensive logging. Nonetheless, it has become the most dominant species in this region. The high abundance of sugar maple as secondary species recorded in the GLO data set also indicates that sugar maple would become a dominant species a hundred years later. The removal of pine and shade-tolerant species such as hemlock and sugar maple and the fragmentation on the landscape due to human activities contributed to the increase of mid-shade-tolerant species such as red oak and basswood, and shade-intolerant species such as aspen and paper birch. These forest types are common in the current landscapes compared with the presettlement landscapes. With the constructed presettlement forest landscape, it is possible to perform the above comparisons in a spatially explicit manner (Radeloff *et al.*, 1998, 1999) and at multiple scales. Furthermore, using GIS, presettlement landscape can be presented using different classifications that are comparable with existing data.

CONCLUSIONS

We have analysed witness tree records from 1839 to 1866 from the GLO survey, using GIS at multiple scales. The results quantitatively described the possible outcomes when GLO records are processed under a given window size. Small window sizes such as the 1×1 section preserve spatial resolutions but derive the most conservative estimates of species percentage area, which increases at larger window sizes. Such increases, under the 2×2-section resolution, are in the order of three to four times for the least common species, two to three times for the medium common species, and one to two times for the most common species. The species' spatial distribution patterns in the landscape strongly affect the order of magnitude of species percentage area changes moving between different scales. Aggregated distributions tend to have small to moderate increases when moving to a coarser resolution, as shown for hemlock and sugar maple. The GIS processing approach computes vegetation indices such as *RIV* for all species, while preserving the original information. *RIV*s derived from this study provide a quantitative basis for comparing species importance and deriving vegetation classification schemes. Analysis of presettlement landscapes can be done at single, as well as multiple, species level. The feasibility of constructing presettlement landscape maps using various classification schemes makes the GLO data set more useful in assisting various practical needs, such as forest management planning and ecosystem restoration at landscape scales.

ACKNOWLEDGMENTS

The project was supported in part by Wisconsin Department of Natural Resources and U.S.G.S Biological Resources Division. We thank the students who assisted with GLO data entry. We appreciate the helpful comments from Steve Ventura and Volker Radeloff.

REFERENCES

Abrams, M. O. & Ruffner, C. M. (1995) Physiographic analysis of witness-tree distribution (1765–1798) and present forest cover through north central Pennsylvania. *Canadian Journal of Forest Research*, **25**, 659–668.

Anderson, R. C. (1970) Prairies in the Prairie State. *Transcripts of the Illinois State Academy of Science*, **63**, 214–221.

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

Brown, D. G. (1998b) Classification and boundary vagueness in mapping presettlement forest types. *International Journal of Geographical Information Systems*, **12**, 105–129.

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

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

Curtis, J. T. (1959) *The Vegetation of Wisconsin: an Ordination of Plant Communities*, University of Wisconsin Press, Madison, WI.

Delcourt, H. R. & Delcourt, P. A. (1996) Presettlement landscape heterogeneity: evaluating grain of resolution using General Land Office Survey data. *Landscape Ecology*, **11**, 363–381.

ESRI, Inc. (1996) *Arc/Info surface modeling with TIN*. ESRI Inc., Redland, CA.

Finley, R. W. (1976) *Original vegetation cover of Wisconsin* (map). North Central Forest Experiment Station, USDA Forest Service, St. Paul, MN.

Fralish, J. S., Crooks, F. B., Chambers, J. L. & Harty, F. M. (1991) Comparison of presettlement second-growth and old-growth forest on six site types in the Illinois Shawnee Hills. *American Midland Nature*, **125**, 294–309.

Galatowitsch, S. M. (1990) Using the original land survey notes to reconstruct presettlement landscapes in the American west. *Great Basin Nature*, **50**, 181–191.

Gordon, R. B. (1969) The natural vegetation of Ohio in pioneer days. *Bulletin of Ohio Biological Survey*, **3**, 1–113.

Hansen, M. H., Frieswyk, T., Glover, J. F. & Kelly, J. F. (1992) *The Eastwide Forest Inventory Database: User's Manual*. USDA Forest Service General Technical Report, NC 151.

He, H. S., DeZonia, B. & Mladenoff, D. J. (in press) An aggregation index to quantify spatial patterns of landscapes. *Landscape Ecology*, in press.

He, H. S., Mladenoff, D. J., Radeloff, V. C. & Crow, T. R. (1998) Integration of GIS data and classified satellite imagery for regional forest assessment. *Ecological Applications*, **8**, 1072–1083.

He, H. S. & Ventura, S. J. (1995) Apostle Islands vegetation coverage—a GIS framework. *Symposium of International Society for Photogrammetry and Remote Sensing*, pp. 59–68. Commission IV, Mapping and Geographical Information Systems, Madison, WI.

Host, G. E., Polzer, P. L., Mladenoff, D. J. & Crow, T. R. (1996) A quantitative approach to developing regional ecosystem classifications. *Ecological Applications*, **6**, 608–618.

Iverson, L. R. (1988) Land-use changes in Illinois, USA: the influence of landscape attributes on current and historic land use. *Landscape Ecology*, **2**, 45–61.

Kapp, P. O. (1978) Presettlement forest of Pine River watershed (Central Michigan) based on original land survey records. *Michigan Botany*, **17**, 3–15.

Keys, J. Jr, Carpenter, C., Hooks, S., Koenig, F., McNab, W. H., Russell, W. & Smith, M. L. (1995) *Ecological units of the eastern United States—first approximation* (map and booklet of map unit tables). USDA Forest Service, Atlanta, GA.

Kunin, W. E. (1998) Extrapolating species abundance across spatial scales. *Science*, **284**, 1513–1515.

Leitner, L. A., Dunn, C. P., Guntenspergen, G., Stearns, R. F. & Sharpe, D. M. (1991) Effects of site, landscape features, and fire regime on vegetation patterns in presettlement southern Wisconsin. *Landscape Ecology*, **5**, 203–217.

Lorimer, C. G. (1977) The presettlement forest and natural disturbance cycle of northeastern Maine. *Ecology*, **58**, 139–148.

Manies, K. L. (1997) *Evaluation of general land office survey records for analysis of the northern Great Lakes hemlock-hardwood forests*, MSc Thesis, University of Wisconsin-Madison, Madison, WI.

Mladenoff, D. J., Dahir, S., Sickley, A. T. & Guntenspergen, G. G. (1999) Probabilistic classification of ambiguously identified tree species in the US General Land Office survey data (Abstract), pp. 23. *Proceedings*, International Association of Landscape Ecology, Snowmass, CO.

Mladenoff, D. J. & Howell, E. A. (1980) Vegetation change on the Gogebic Iron Range (Iron County, Wisconsin) from the 1960s to present. *Transcripts of the Wisconsin Academy of Science, Arts and Letters*, **68**, 74–89.

Mladenoff, D. J. & Pastor, J. (1993) Sustainable forest ecosystems in the northern hardwood and conifer forest region:

concepts and management. *Defining sustainable forestry* (ed. by G. H. Aplet, N. Johnson, J. T. Olson and V. A. Sample), pp. 145–180.

Nelson, J. C. (1997) Presettlement vegetation patterns along the 5th Principal Meridian, Missouri territory, 1815. *American Midland Nature*, **137**, 79–94.

Nelson, J. C., Redmond, A. & Sparks, R. E. (1994) Impacts of settlement on floodplain vegetation at the confluence of the Illinois and Mississippi Rivers. *Transcripts of the Illinois State Academy of Science*, **87**, 117–133.

Radeloff, V. C., Mladenoff, D. J., He, H. S. & Boyce, M. S. (1999) Forest landscape change: the northwest Wisconsin Pine Barrens before European settlement and today. *Canadian Journal of Forest Research*, **29**, 1649–1659.

Radeloff, V. C., Mladenoff, D. J., Manies, K. L. & Boyce, M. S. (1998) Pre-settlement and current composition of oak in the northwest Wisconsin Pine Barrens landscape. *Transcripts of the Wisconsin Academy of Science, Arts and Letters*, **86**, 189–205.

White, M. A. & Mladenoff, D. J. (1994) Old-growth forest landscape transitions from pre-European settlement to present. *Landscape Ecology*, **9**, 191–205.

Whitney, G. G. (1982) Vegetation-site relationships in the pre-settlement forests of northeastern Ohio. *Botany Gazette*, **143**, 225–237.

Whitney, G. G. (1990) The history and status of the hemlock-hardwood forests of the Allegheny Plateau. *Journal of Ecology*, **78**, 443–458.

BIOSKETCHES

Hong S. He is an assistant professor at the University of Missouri-Columbia. **David J. Mladenoff** is a forest ecosystems and landscape ecologist and associate professor, and **Theodore A. Sickley** is a research specialist, all in the Department of Forest Ecology and Management, University of Wisconsin-Madison. **Glenn R. Guntenspergen** is a landscape ecologist with the United States Geological Survey. Dr He primarily works in the field of landscape ecology and geographical information systems, focussing on spatially explicit landscape modelling, GIS and remote sensing data integration, GIS interpolation and analytical approaches. Dr Mladenoff is interested in forest landscape dynamics, particularly involving the interactions of natural disturbances with human use of landscapes. Mr Sickley works on natural resources applications of GIS as well as database development. Dr Guntenspergen is interested in the link between spatial phenomena and the processes regulating plant community structure and organization.