Journal of Biogeography (*J. Biogeogr.*) (2007) **34**, 500–513

ORIGINAL ARTICLE

Spatial patterns and vegetation–site relationships of the presettlement forests in western New York, USA

Yi-Chen Wang*

Department of Geography, University at Buffalo, The State University of New York, Buffalo, NY 14261, USA

*Correspondence: Yi-Chen Wang, Department of Geography, National University of Singapore, 1 Arts Link, Singapore, 117570.
E-mail: geowyc@nus.edu.sg

ABSTRACT

Aim The purposes of this study were to develop a Geographic Information System and spatial analytical methodology to reconstruct and represent the presettlement vegetation in a spatially continuous manner over large areas and to investigate vegetation–site relationships before widespread changes of the vegetation had taken place.

Location The study area was the Holland Land Company Purchase in western New York, a 14,400 km^2 area extending across the physiographic provinces of the Erie–Ontario Lowlands and the Appalachian Uplands.

Methods Bearing-tree records from the Holland Land Company township surveys of western New York in *c.* 1800 were collected and analysed. The geostatistical method of indicator kriging was used to map spatially continuous representations of individual tree species. Rule-based and statistically clustered approaches were used to analyse and classify the reconstructed tree species distributions in order to obtain the vegetation association distribution. Contingency table analysis was conducted to quantify species relationships with soil conditions.

Results The presettlement vegetation at both the tree species and the vegetation association levels were easier to interpret and visually more effective as a spatially continuous representation than as a discontinuous distribution of symbols. The results for tree species were probability occurrences of species distribution, showing spatial patterns that were not apparent in discrete maps of points or in summary tables of species frequencies. Analysis of the 8792 bearing trees suggested the dominance of American beech (*Fagus grandifolia*) and sugar maple (*Acer saccharum*) in the forest composition 200 years ago. Both soil drainage and texture were important site determinants of the vegetation in western New York. The rule-based and statistically clustered approaches had the advantage of summarizing vegetation compositional patterns in a single image, thus avoiding the need to delineate manually and subjectively the location of boundaries between adjacent vegetation associations.

Main conclusions The study offers more insights into the spatial pattern of presettlement forests in western New York than do prior studies. The spatially continuous representation could also enable the comparison of vegetation distribution from data sources that have different sampling schemes, for example the comparison of presettlement vegetation from the presettlement land survey records with current vegetation from modern forest inventories. The results are of value, providing a useful benchmark against which to examine vegetation change and the impacts of human land use.

doi:10.1111/j.1365-2699.2006.01614.x

Keywords

Bearing trees, geostatistics, Holland Land Company, land survey records, presettlement forest, vegetation mapping, vegetation reconstruction, vegetation–site relationship, western New York.

INTRODUCTION

Ecological biogeographers typically study ecosystems using one of three foci: (1) the mechanisms through which organisms and populations interact with each other and their environment; (2) the temporal dynamics and spatial patterns of species and communities that result from those interactions; and (3) the ways in which human activities can negatively impact natural systems in the form of global change, and the ways in which humans counteract those negative impacts through the positive practices of ecosystem management (Cox & Moore, 1993; MacDonald, 2003). These foci are, of course, interrelated. An understanding of how much change has been induced by humans requires reconstruction of the states of ecosystems at different time periods (Schulte *et al.*, 2002). Similarly, the development of successful and sustainable ecosystem restoration and management requires knowledge of the pre-impacted state of ecosystems so as to match species with their appropriate site requirements (Egan & Howell, 2001). Finally, insights into the physical and biological mechanisms by which organisms interact allow the most destructive human practices to be identified (Dale & Haeuber, 2001).

In this paper I work within the second ecological–biogeographical focus of reconstructing the spatial patterns of plant species and communities prior to major human impact. Various sources of information are available to reconstruct historical vegetation conditions, such as pollen records, presettlement land survey records (PLSRs), and travellers–accounts (Egan & Howell, 2001). The PLSRs are of particular interest to many biogeographers and ecologists in North America because the data were collected just before European settlement and the attendant large-scale reorganization of the landscape. Analyses of the PLSRs have been used to reconstruct various characteristics of the pre-European settlement vegetation of an area: percentage abundance and importance values of individual species (Wuenscher & Valiunas, 1967; Delcourt & Delcourt, 1974; Nelson, 1997); types of vegetation communities (Veatch, 1925; Schulte *et al.*, 2002); and type and rate of disturbance (Lorimer, 1977; Schulte & Mladenoff, 2005). These characteristics have been expressed as summary statistics for a complete study area (Spurr, 1951; Shanks, 1953) and for different parts of a study area (McIntosh, 1962; Cogbill *et al.*, 2002). These characteristics have also been represented spatially using various cartographic practices.

Most maps of presettlement vegetation patterns have utilized discrete point symbols to express the presence of individual species or the dominant community types (Howell & Kucera, 1956; Grimm, 1984; Leitner *et al.*, 1991; Seischab, 1992). Community types have also been mapped as discrete polygons (Siccama, 1971); however, before the development of Geographic Information Systems (GIS) and spatial analysis techniques, the polygons were manually and subjectively placed using apparently homogeneous patterns of abundant species.

The GIS and spatial techniques that are now available allow spatially continuous representations of PLSRs to be created using objective quantitative analyses (Brown, 1998a; Wang and Larsen, in press). These spatially continuous representations are preferable to spatially discrete representations for at least three reasons. First, spatial continuity of species distributions are more in keeping with the continuum concept in ecology (Gleason, 1926; Whittaker, 1951) than are the discrete representations. Second, the translation of discrete points into continuous surfaces allows the gaps between data points to be filled (Johnston, 1998) and the results to be visualized in effective and meaningful ways (Kemp, 1997) that are easier for users such as environmental managers to interpret. Third, continuous representations enable the comparison of vegetation distributions from data sources that have different sampling schemes. For example, vegetation change information could be obtained by comparing vegetation patterns reconstructed from past PLSRs with those from modern forest inventories, and the results could provide insights into human influences on the landscape.

The focus of this study is the reconstruction of the presettlement vegetation in western New York using a spatially continuous representation that is founded on the continuum concept (Gleason, 1926; Whittaker, 1951) and the field point of view (Goodchild, 1992). Continuous fields are richer portrayals of vegetation distribution than discrete objects because they can illustrate gradual transitions, or ecotones, between vegetation types (Brown, 1998a). Readers seeking background information on the object–field discussions for the representation of geographic phenomena are direct to Goodchild (1992, 1994) or Burrough (1996). The specific objectives of this study are to: (1) reconstruct the presettlement tree species distributions from point bearing-tree data to provide a more continuous, as opposed to discrete, representation; (2) propose a quantitative and replicable approach to analysing the reconstructed species distributions to obtain vegetation associations; and (3) investigate vegetation–site relationships before widespread changes of the vegetation in western New York had taken place. The study will improve the understanding of historical vegetation patterns and their relationships to site conditions, which may be useful for further ecosystem management.

STUDY AREA AND DATA

Study area

The study area of western New York lies between Pennsylvania and Lake Ontario and is bordered on the west by Lake Erie and the Niagara River. This 14,400 km^2 area, between 78° and 79°30′ W, was purchased and surveyed by the Holland Land Company (HLC).

Western New York has a humid continental climate with a mean annual temperature of 8.3°C and a mean annual precipitation ranging from 800 mm in the north to 1120 mm in the south (Easterling *et al.*, 1996). The study area extends across two recognized physiographic provinces (Fig. 1a). The northern part of the study area is located in the Erie–Ontario Lowlands, a province with relatively low, flat topography that has been substantially modified by the glacial deposition of moraines, drumlin fields, and shoreline deposits. The southern part belongs to the Appalachian Uplands and has a complex topography with generally thin glacial deposit cover. Only the Allegany State Reservation on the western part of the New York–Pennsylvania border escaped glaciation (Broughton, 1966). Braun (1950) recognized two forest types in the study area, and their distributions approximated the two physiographic provinces: beech–maple (*Fagus–Acer*) forests in the Erie–Ontario Lowlands, and hemlock–white pine–northern hardwood forests (*Tsuga–Pinus–Acer–Betula–Fagus*) in the Appalachian Uplands. Küchler's (1964) map of the potential natural vegetation placed most of the Appalachian Uplands in the northern hardwood forests, with extensions of the more southern Appalachian oak (*Quercus*) forests penetrating the major river valleys. The predominant land cover of the study area is second-growth deciduous and coniferous forests.

Holland Land Company survey records

A spatially comprehensive and temporally relevant representation of the presettlement vegetation of western New York is contained in the private Holland Land Company (HLC) survey of *c.* 1800. The survey of the HLC purchase was conducted in two stages. The first stage involved the survey of the lakeshore boundary and township perimeters between 1797 and 1799, a

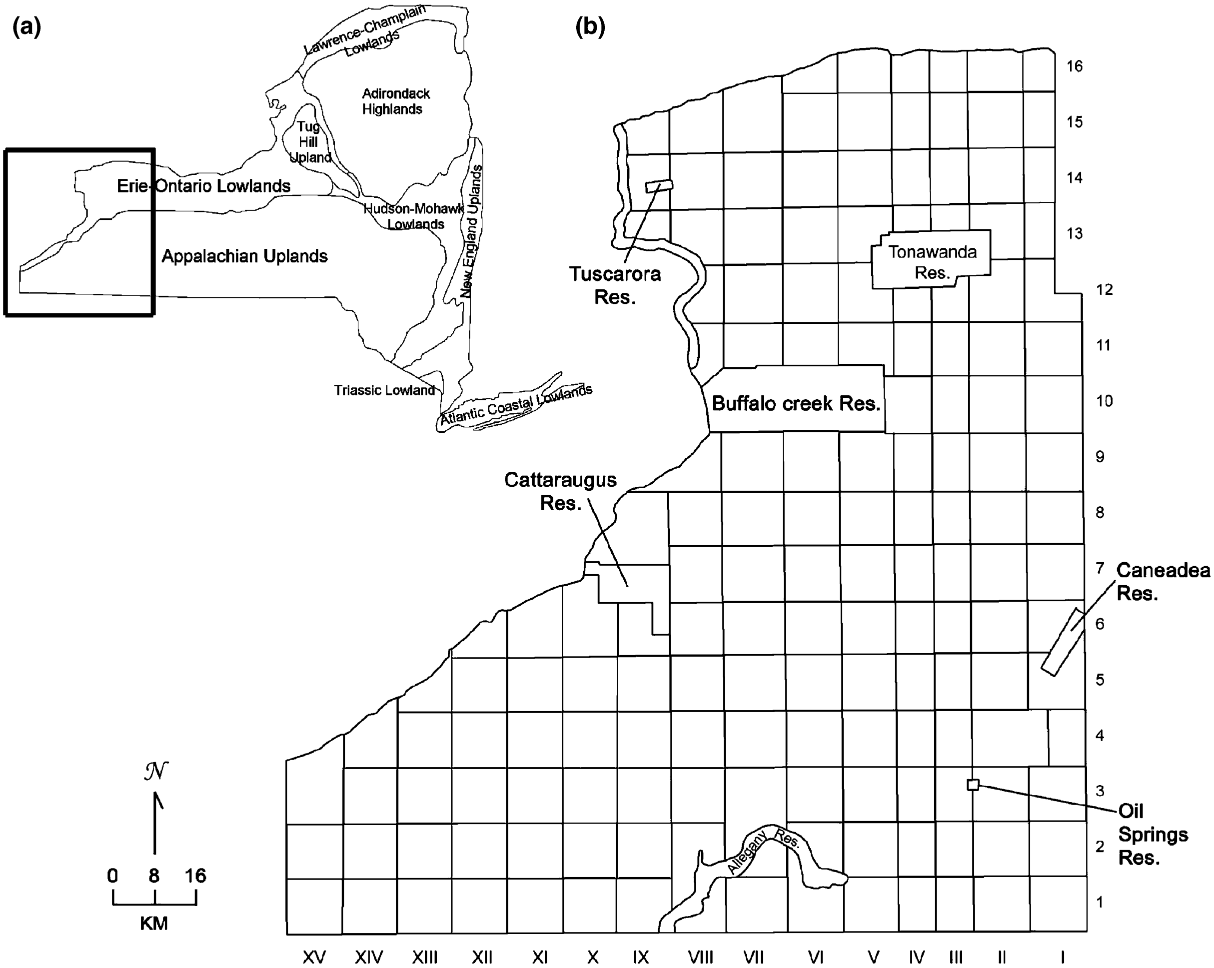

Figure 1 (a) Physiographic provinces of New York based on relief and geology, with the study area outlined [modified from Broughton (1966), p. 32]. (b) The study area of the Holland Land Company Purchase and its township survey outlines in western New York [modified from Wyckoff (1988), pp. 29 and 136].

time period when the land was only thinly populated by scattered bands of Native Americans and was still little known to American settlers. The second stage was the subdivision of townships into lots, mostly conducted between 1799 and 1814 (Wyckoff, 1988). Just as later public General Land Office (GLO) surveys that divided land into townships of 6 × 6 miles (1 mile = 1.6 km), most of the townships of the private HLC surveys had a size of 6 × 6 miles, but 4 × 6 and 7 × 6 miles were also found (Fig. 1b). This resulted in a total of 162 townships. Surveyors set up posts every half-mile along the township perimeters, and recorded two to four nearby bearing trees for the survey posts. The surveyors also recorded line descriptions providing tree species lists and soil quality for several segments of each surveyed mile, as well as sketch maps describing the terrain features that crossed the survey lines.

Surveyors' manuscripts on microfilms were obtained from the HLC archives of the State University of New York College at Fredonia. Bearing-tree data of the township perimeter surveys were used in this study. This is supported by the rationale that the township-level data provide a more temporally precise picture of the presettlement vegetation than the information from the section- or lot-level data because the township perimeter surveys were usually conducted first and accomplished within a few years, as opposed to the time span of several decades of the section or lot subdivision surveys. Moreover, results from Wang (2004) and Wang and Larsen (2006) have indicated that coarsely resolved 6 × 6 mile township-level data (i.e. bearing trees recorded at every half-mile along the 6-mile township lines) provide a spatial pattern of presettlement vegetation over a large area that is similar to that obtained from finely resolved 1 × 1 mile section-level data.

The quality of PLSRs are often of concern regarding imprecise quantifications of distance and bearing and erroneous identification of tree species (Wang, 2005). The HLC surveyors, Joseph Ellicott and his survey crew, were considered more accurate than previous surveyors in New York State because they used transits rather than the less accurate hand-held compasses, and because axemen were employed to remove trees along the line of sight in order to increase the accuracy of the survey (Seischab, 1992). The reputation of the survey crew was good, suggesting that fraudulent surveys should not be an issue and that the identification of tree species was probably good (Wyckoff, 1988).

METHODS

Data collection and process

Information on survey corners and their associated bearing trees were transcribed into MS Excel spreadsheets from microfilms. It was not possible to use a Public Land Survey digital base map as a template for topological linking and deriving the survey corners and bearing trees in to GIS as was done in other studies (Batek *et al.*, 1999; Dyer, 2001), because a similar digital base map was not available for the private HLC surveys. Instead, the south-west corner of New York State, the starting point of the HLC township survey, was used as the origin. A data conversion program was developed using C++ programming language to obtain the coordinates of the subsequent survey corners and bearing trees based on the recorded distances and bearings. The obtained coordinates were next converted into GIS and further spatially adjusted using the modern town corners that matched the HLC survey corners so as to increase the positional accuracy of the data. The adjusted bearing-tree locations in GIS point coverages were then used in reconstructing presettlement vegetation and investigating vegetation–site relationships. Data processing and analyses were conducted using ArcView and ArcGIS (ESRI Inc., Redlands, CA, USA) unless otherwise specified.

Reconstruction of presettlement tree species distributions

Various methods have been used in PLSR studies to reconstruct presettlement vegetation, for example environmental modelling and spatial interpolation. Wang and Larsen (2006) compared these methods and suggested that, in the context of the PLSR studies in western New York, the most useful approach to deriving a spatially continuous representation for tree species is the geostatistical method of indicator kriging. Although Manies & Mladenoff (2000) considered that indicator kriging was inappropriate for capturing the complexity of presettlement landscape at small extents of less than 100 km^2, they did suggest that the accuracy of the method would be increased over larger extents such as a county or a state, as in this study, which covers 14,400 km^2.

The analytical procedure of Wang & Larsen (2006) was employed here. To perform indicator kriging, the original categorical values of tree species names were transformed to binary values of 1 and 0. The value of 1 indicated the presence of a specific species, and 0 indicated the absence of that species but the presence of some other species at that point location. Semi-variograms for individual species were calculated using the transformed binary data and then visually fitted with semi-variogram models that were applied for interpolation at unsurveyed locations to predict the probability of a species' occurrence. The output map of indicator kriging portrayed the spatial distribution for each species as a continuous probability surface of values ranging from zero to one. The probability surfaces were then used as input data to reconstruct presettlement vegetation associations.

Reconstruction of presettlement vegetation associations

Classification methods were developed to represent compositional variation among plant assemblages and obtain vegetation associations for three reasons. First, regardless of whether or not vegetation communities are real entities, some form of classification is a practical necessity for mapping purposes (Whittaker, 1970; Küchler & Zonneveld, 1988; Brown, 1998b). Second, discontinuities in the environment can lead to discontinuities in communities (Forman, 1995). Third, the interactions of species with each other result in certain combinations of species tending to recur together in certain environments (Austin & Smith, 1989). In summary, relative discontinuities in the continuum allow the partition of a continuous landscape into reasonably discrete ecological units (Bryer *et al.*, 2000).

The presettlement distribution of vegetation associations was created using the continuous probability surfaces of individual species from indicator kriging. The probability surface for each species was resampled into 1 × 1 mile grid cells – a GIS raster data structure of continuous fields. Only taxa that made up at least 1% of the bearing-tree data base (He *et al.*, 2000) were used to obtain vegetation associations. The grids were classified using two methods: a rule-based approach that employed the raw probability data, and a statistically clustered approach that employed normalized probability data.

Rule-based approach to association classification

Rule-based approaches to classification use rules and numerical thresholds to interpret information represented in multiple data layers (Johnston, 1998). In the case of vegetation association reconstruction, the multiple data layers were the resampled grids of probability surfaces of different species. The resampled grids of different species were overlaid and classified based on all species modelled as present with a probability greater than 0.3, a cutoff value suggested by Batek *et al.* (1999) and Manies & Mladenoff (2000) to avoid using data from cells with extremely low probabilities. Vegetation associations were defined solely based on joint occurrence to avoid grouping them by contemporary vegetation types that may not have existed 200 years ago (Batek *et al.*, 1999).

Statistically clustered approach to association classification

Cluster analysis is an exploratory data analysis tool that sorts cases (here, individual cells with different species' probabilities) into groups. Clustering was performed in three steps. The first step involved creating a 'cell template' that was actually a polygon coverage consisting of a matrix of 1 × 1 mile squares. The second step was to compile the resampled grids of probability surfaces of different species and to normalize the probabilities within a cell so that, for each cell, the sum of all species' probability occurrences became one. The normalized probabilities were then assigned to the designated polygon square of the new cell template. The third step involved importing the attribute table associated with the polygon coverage of a matrix of squares into SPSS software (SPSS Inc., Chicago, IL, USA) for cluster analysis. The Euclidean distance measure was used along with Ward's method to obtain clusters. The number of clusters was chosen by trying different numbers so that there were meaningful distinctions between groups but not so many that they became overwhelming or confusing (Campbell, 2001).

Investigation of vegetation–site relationships

The site relationships of bearing-tree species were investigated with respect to soil texture and drainage. These soil properties are most important to plant growth (Curtis, 1959; Barrett *et al.*, 1995) and are available in the two commonly used digital soil data bases from the USDA Natural Resources Conservation Service: the State Soil Geographic (STATSGO) data base and the Soil Survey Geographic (SSURGO) data base. In this study, the STATSGO data base of New York was used for two reasons, even though it was more coarsely resolved than the SSURGO data base. First, the STATSGO data base is appropriate for analysis at the multi-county level (USDA NRCS, 1995) and was used in examining presettlement vegetation–site relationships in east-central Alabama by Black *et al.* (2002). The size of their study area is similar to that of the HLC land. Second, the SSURGO data base does not provide both spatial and attribute data for the whole study area.

The investigation was performed in three steps. First, all the components for each map unit (polygon) in the STATSGO data base were combined to derive the dominant soil-surface texture and drainage. Second, the bearing-tree points were overlaid with soil polygons, and all trees within a buffer zone of a soil boundary were eliminated to increase the certainty of locating the trees on the correct soil series. The STATSGO data base provided a coverage at a scale of 1 : 250,000; following the procedure of Barrett *et al.* (1995), a 125-m buffer was chosen. Third, contingency table analysis was conducted to quantify associations of species and soil properties (Strahler, 1978) for the taxa that accounted for at least 2% (Dyer, 2001) of the remaining trees after buffering in order to minimize errors associated with small sample sizes (Sokal & Rohlf, 1995). For those species with a significant G statistic, frequency counts in the contingency tables were converted into standardized residuals according to the method of Haberman (1973). The residuals quantified the degree of preference or avoidance of a species for a particular soil condition.

RESULTS

Reconstruction at the tree species level

Although the identification of tree species may be considered good, given the HLC survey crew's reputation (Wyckoff, 1988), it was common for multiple names to be used for the same species. More than 40 tree types were mentioned by the surveyors. Some of the tree types used by the surveyors actually referred to the same species, for example basswood (*Tilia americana*) and lyndon. A total of 38 taxa were finally identified, excluding the name of 'Tree', which was unclear in its meaning and was used by the surveyors only three times (Table 1). Common names used by the surveyors were generally the same as those used today. Collective names such as birch, oak, and pine used in the survey records lead to some ambiguity of species names. Thus, assumptions were made in the interpretation of the common names. For example, both Seischab (1992) and Whitney & DeCant (2001) suggested that birch was predominately yellow birch (*Betula alleghaniensis*), although it might have included a few black birch (*B. lenta*), and hence in this study birch was interpreted as yellow birch.

Table 1 Presettlement forest composition in western New York. The column of taxa shows surveyor designations. The frequencies include all the bearing trees (diameter cutoff not specified by surveyors) recorded in the HLC notes.

Taxa	Taxonomic equivalents	Frequency	Percentage
Beech	*Fagus grandifolia*	3256*	37.0
Sugar maple	*Acer saccharum*	1848*	21.0
Hemlock	*Tsuga canadensis*	731	8.3
Basswood	*Tilia americana*	450*	5.1
Elm	*Ulmus* spp.	400	4.5
Black ash, ash	*Fraxinus nigra*	330	3.8
White oak, oak	*Quercus alba*	259	2.9
Red maple	*Acer rubrum*	225*	2.6
White ash	*Fraxinus americana*	199	2.3
Yellow birch, birch	*Betula alleghaniensis*	215†	2.4
White pine, pine	*Pinus strobus*	178	2.0
Chestnut	*Castanea dentata*	105	1.2
Ironwood	*Ostrya virginiana*	105*	1.2
Black oak	*Quercus velutina*	99	1.1
Hickory	*Carya* spp.	95	1.1
Cucumber	*Magnolia acuminata*	55	0.6
Poplar, aspen, aspine	*Populus* spp.	41	0.5
Cherry	*Prunus serotina*	40	0.5
Butternut	*Juglans cinerea*	28	0.3
Tamarack	*Larix laricina*	25	0.3
Sycamore	*Platanus occidentalis*	24	0.3
Cedar	*Thuja occidentalis*	11	0.1
Red oak	*Quercus rubra*	11	0.1
Black birch	*Betula lenta*	7	0.1
Black walnut	*Juglans nigra*	7	0.1
Willow	*Salix nigra*	6	0.1
Fir	*Abies balsamea*	5	0.1
Gum	*Liquidambar styraciflua*	5	0.1
Swamp oak	*Quercus palustris*	5	0.1
Chestnut oak	*Quercus prinus*	5	0.1
Hornbean	*Carpinus caroliniana*	4	‡
Whitewood	*Liriodendron tulipifera*	4	‡
Rock oak	*Quercus montana*	3	‡
Tree	–	3	‡
Spruce	*Picea mariana*	2	‡
Swamp white oak	*Quercus bicolor*	2	‡
Thorn	*Crataegus* spp.	2	‡
Alder	*Alnus incana*	1	‡
Red elm	*Ulmus rubra*	1	‡
Total		8792	100

*Including one sapling.

†Including two saplings.

‡Less than 0.1%.

In most cases, the interpretations by Seischab (1992) were followed because the study area and surveyors were the same as for this study, even though quantitative point bearing-tree data were used here while Seischab (1992) used qualitative survey-line descriptions. For example, in Seischab (1992), the surveyors' designations for poplar, aspen, and aspine were interpreted as *Populus* spp., and whitewood and tulip trees were interpreted as yellow poplar (*Liriodendron tulipifera*). One major difference, however, was that in Seischab's study, maple was interpreted to mean sugar maple (*Acer saccharum*), red maple (*A. rubrum*), or silver maple (*A. saccharinum*), according to its association with different species in the line descriptions. In this study maple was taken to mean red maple but not sugar maple, because it was found that surveyors used both maple and sugar maple to record bearing trees at the same townships and even at the same survey corners. It is therefore believed that the surveyors did differentiate the two species when recording bearing trees. This assumption is similar to that used in Whitney & DeCant (2001).

Relative frequency

A total of 8792 bearing trees were recorded in the HLC township surveys (Table 1), seven of which were recorded as 'saplings'. Of the 38 taxa involved, 15 made up at least 1% of the bearing-tree records. The two most abundant species were American beech (*Fagus grandifolia*) and sugar maple, accounting for, respectively, 37% and 21% of the total bearing trees. Other abundant species were eastern hemlock (*Tsuga canadensis*), basswood, and elm (*Ulmus* spp.). These five taxa together accounted for about 76% of the bearing trees. As a group, ash made up 6% of the sample, while oak made up about 4.3%, with white oak at 2.9% and black oak at 1.1%. American chestnut (*Castanea dentata*) constituted about 1.2% of the presettlement forest.

Spatial distributions of common species

Indicator kriging was conducted for the 15 most common taxa to create their spatial distributions. In addition, the spatial distribution of tamarack (*Larix laricina*) was recreated and included in the following association-level reconstruction because, even though this species constituted only 0.3% of the total bearing trees, its concentrated distribution and strong spatial autocorrelation make it an important component of the landscape pattern.

Spatial distributions of beech, sugar maple, hemlock, black ash (*Fraxinus nigra*), white pine (*Pinus strobus*), and tamarack, six taxa that are either common or have unique spatial distributions, are shown as non-kriged point maps of bearing-tree locations and continuous surfaces of predicted probability occurrences reconstructed from indicator kriging (Fig. 2).

Beech occurred throughout the study area (Fig. 2a) but was notably absent from the areas within the Erie–Ontario Lowland province close to its border with the Appalachian Upland province. Sugar maple had its core distribution located in the east-central portion of the study area (Fig. 2b). Hemlock frequently occurred in the region located to the south of Lake Erie and in the southern part of the study area, along the New York–Pennsylvania state border (Fig. 2c). The distribution of this species was clearly negatively associated with the distribution of sugar maple (Fig. 2b). The distribution of black ash was primarily within the Erie–Ontario Lowland province (Fig. 2d), the northern half of the study area. White pine was primarily found in the Caneadea Reservation and in the

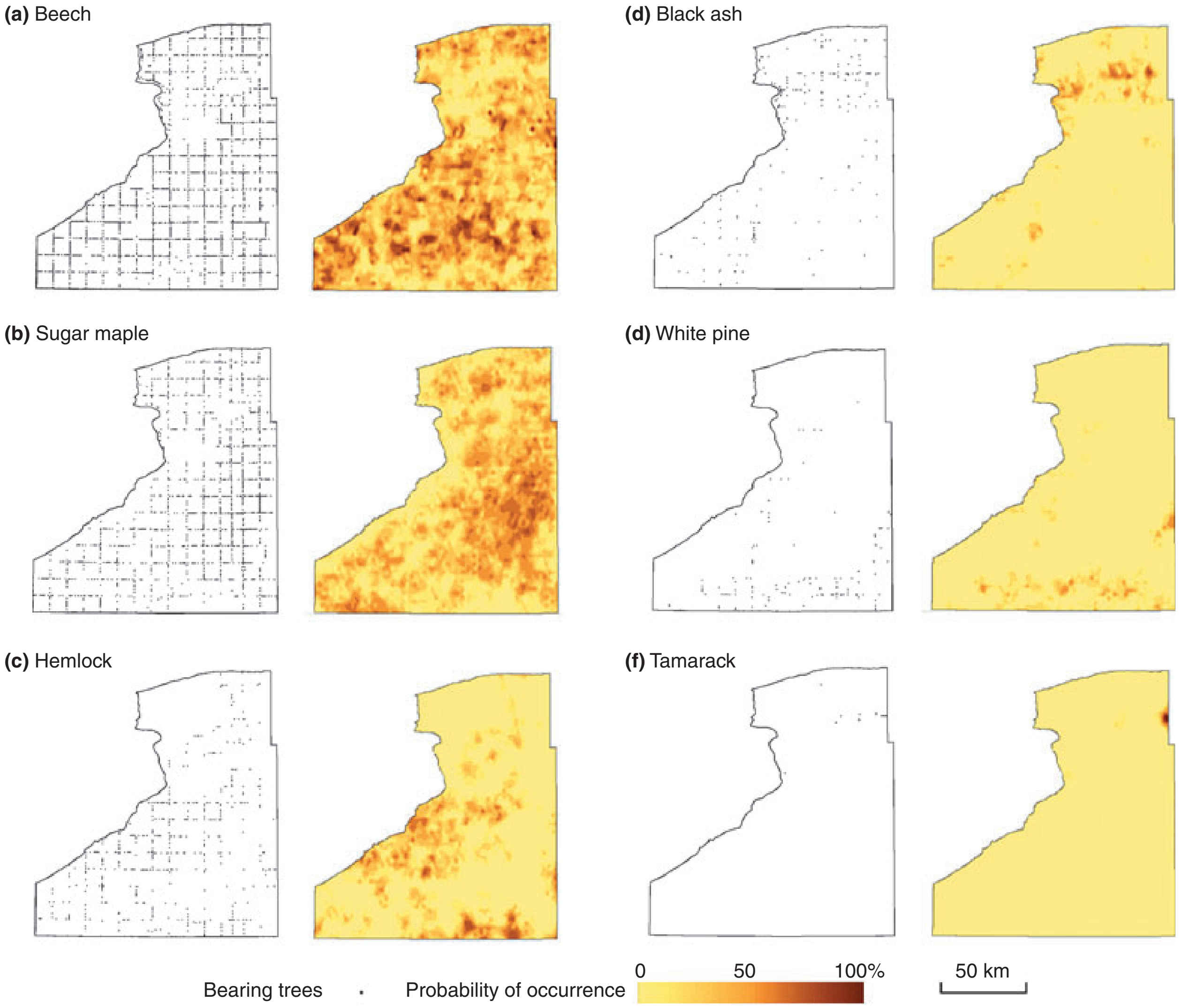

Figure 2 Discrete point distributions and continuous probability surfaces for selected taxa.

southern townships near the Allegany Reservation (Fig. 2e). Tamarack exhibited marked dominance in a small portion in the north-east of the study area, within Township No. 14 of the First Range (Fig. 2f).

Reconstruction at the vegetation association level

Association classifications were based on 16 taxa: the 15 taxa that each constituted more than 1% of the bearing-tree data base (Table 1), and tamarack, owing to its high probability of occurrence within a small area.

Rule-based reconstruction

A total of 212 association types were indicated by the various species combinations that met the 0.3 joint probability threshold. The name given to each association type included all of the species that met the 0.3 probability level. No weights were assigned to species with higher probabilities (Batek *et al.*, 1999), and so cells classified as 'beech–sugar maple' could mean that either beech or sugar maple had a higher probability of occurrence than the other species. Fourteen association types occurred over at least 1% of the study area, together covering 77.8% of the area (Table 2). The other 198 association types were minor associations, which individually covered less than 1% of the study area, and collectively covered 15.9% of the area. Among these minor associations, four association types were shown on the map because they contained either tamarack or chestnut; the others were unnamed. A total of 6.3% of the study area had no species that met the 0.3 probability threshold, and so did not belong to any association type.

The association types created using the rule-based approach exhibited a high degree of spatial interspersion, but five aggregations were evident (Fig. 3a). First, in the northern part of the study area there was an east–west-trending aggregate of beech-dominated forests. Second, south of this was an aggregate of black ash, interspersed with elm and tamarack associations. Third, a large region in the east-central part of the study area was dominated by sugar maple and beech–sugar maple associations. Fourth, there was a very large region

Table 2 Presettlement vegetation associations of western New York reconstructed using the rule-based approach.

Association	Percentage of study area
Beech	34.9
Beech–sugar maple	11.6
Sugar maple	7.7
White pine–others	3.3
White oak	2.9
White oak–beech	2.5
Beech–elm	2.4
Beech–yellow birch	2.4
Hemlock–beech	2.3
Hemlock	1.7
Black ash	1.7
Elm	1.4
Sugar maple–elm	1.2
Beech–black ash	1.0
Chestnut	0.4
Tamarack	0.2
Tamarack–elm	0.1
Chestnut–beech	0.1
Minor association < 1% of the study area	10.9
Minor association < 0.05% of the study area	5.0
No species > 0.3 probabilities	6.3
Total	100.0

Table 3 Presettlement vegetation associations of western New York reconstructed using the statistically clustered approach.

Association	Percentage of study area
Beech–maple	18.9
Beech	15.7
Sugar maple-rich mesic	15.2
Hemlock–northern hardwood	14.5
Maple–basswood	8.6
Oak forest	8.2
Oak–white pine–chestnut	7.4
Elm–ash	6.0
Black ash–elm	5.1
Tamarack–cedar	0.4
Total	100.0

dominated by beech that ranged from the south-west over to the east, surrounding the third region, and interspersed with small areas of many different associations. The fifth aggregation was the Allegany Reservation area in the south, consisting of hemlock, chestnut, white oak, and white pine, with other minor representations. The unnamed minor associations were distributed randomly throughout the study area, with some aggregations in the northern region of the study area, mainly to the south of the black ash, elm, and tamarack aggregates. The areas in which no species reached the 0.3 probability threshold, indicated by white on the map (Fig. 3a), were most common in transitional areas between the large beech-dominated cluster and adjacent aggregates.

Statistically clustered reconstruction

Cluster analysis classified the cells into units with similar species occurrence probabilities. The cells were clustered into

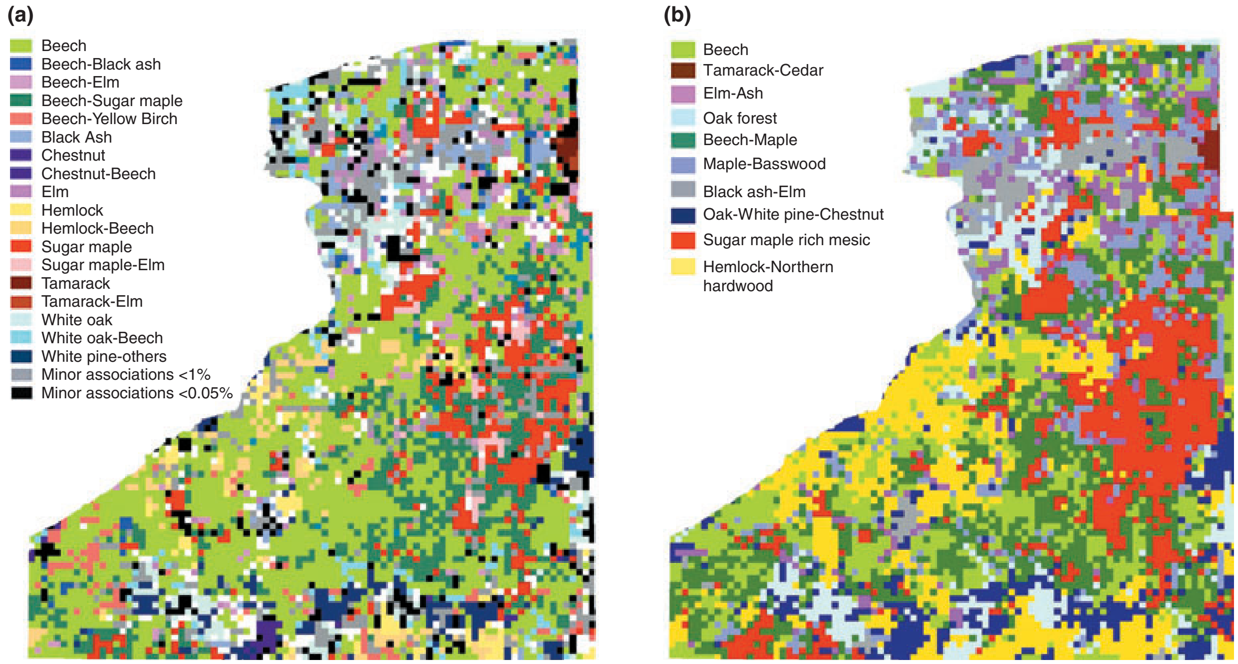

Figure 3 Presettlement vegetation association distributions reconstructed using (a) the rule-based approach and (b) the statistically clustered approach. Note that the white cells in (a) indicate that no species' probability was greater than the cutoff value of 0.3. The resolution for both maps (each square) is 1 × 1 mile.

10 types to allow key patterns to be revealed but without including too many classes.

Each statistical cluster was named (Table 3; Fig. 3b) using a combination of the taxa that had the highest absolute mean probabilities in that cluster as compared with other taxa in that cluster, and the taxa that had their highest relative mean probabilities in that cluster relative to other clusters. For example, if black ash and elm had higher probabilities in most cells of a unit than other units, then the unit was assigned as a black ash–elm association. The names of some associations were modified slightly to correspond to the ecological communities identified and described, but not mapped, by the New York State Department of Environmental Conservation (Edinger *et al.*, 2002). For example, the community called 'sugar maple-rich mesic' was created by Edinger *et al.* (2002), and was used here because its description matched the composition and spatial location of cells that were statistically clustered together in this study.

The two most common associations, beech and beech–maple, were interspersed and most common in the southern portion of the study area (Fig. 3b). They represented hardwood forests where beech dominated and beech and sugar maple co-dominated, respectively. The sugar maple-rich mesic association, a very distinct aggregate in the east-central region, was dominated by sugar maple with associate tree species of yellow birch and beech. The hemlock–northern hardwood association formed an extensive distribution along Lake Erie and the border of New York and Pennsylvania states. In any one cell, hemlock could form nearly pure stands or be co-dominant with the following species: beech, sugar maple, red maple, white pine, yellow birch, red oak (*Quercus rubra*), and basswood. The oak forest association in the southern part of the Appalachian Upland was dominated by white oak, while in the Erie–Ontario Lowland it was composed of black oak (*Quercus velutina*), red oak, white oak, and hickory (*Carya* spp.), an oak–hickory association. The remainder of the associations either formed east–west-trending aggregates, including the black ash–elm association in the north with the tamarack–cedar association at the east side of the band, and the oak–white pine–chestnut association in the south, or were

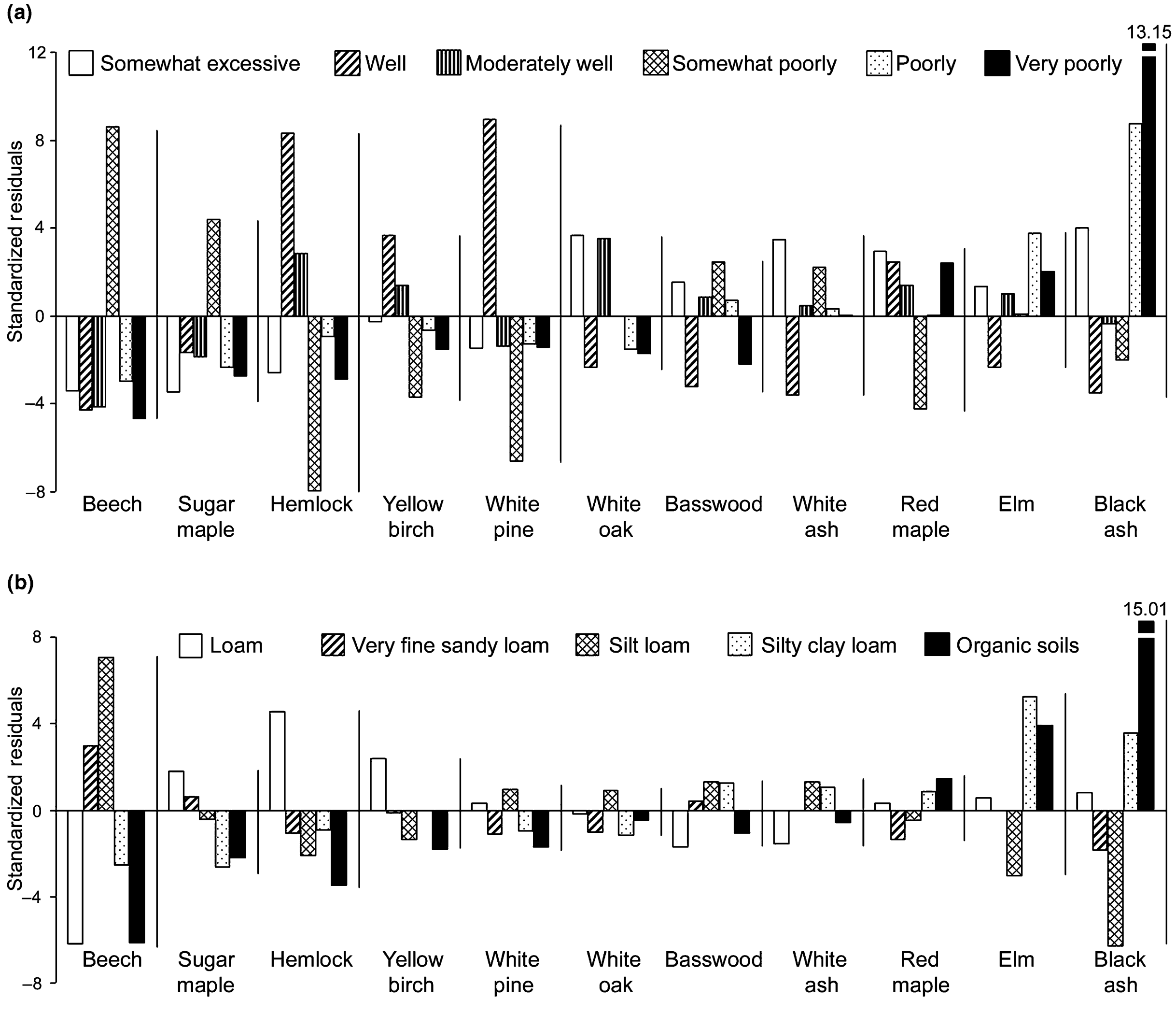

Figure 4 Response of taxa accounting for ≥ 2% of bearing trees to (a) soil drainage and (b) soil texture. Positive or negative residuals (Haberman, 1973) indicate a preference or avoidance, respectively, for the soil condition.

scattered in small aggregates, such as the maple–basswood and the elm–ash types.

Relationships of vegetation-site conditions

Both soil drainage and texture were important determinants of the vegetation in western New York, but the responses of the various taxa to drainage were more apparent than those to texture, as seen in the greater absolute residual values (Fig. 4). In terms of drainage, black ash and elm were strongly associated with poorly and very poorly drained soils. Although beech and sugar maple were characteristic species of somewhat poorly drained soils, they rarely occurred on poorly and very poorly drained sites, a pattern similar to that previously observed in northern Ohio (Whitney, 1982; Whitney & Steiger, 1985). Compared with beech and sugar maple, red maple demonstrated a completely opposite drainage preference and avoidance. In addition to the very poorly drained sites, red maple occurred on soils from moderately well drained to somewhat excessively drained (Fig. 4a). Taxa that preferentially occurred on moderately well drained sites were hemlock, white pine, and yellow birch.

In terms of texture, tamarack, black ash, elm, and red maple showed similar relationships with specific soil conditions: they were positively associated with silt clay loam and organic soils but negatively associated with silt loam. Sugar maple, hemlock, and yellow birch preferred a loamy soil texture, for which beech showed a strong negative association. On the other hand, beech was more frequently observed on very fine sandy loam and silt loam. White pine, white oak, and white ash also tended to occur on soils of silt loam. Their degrees of association for texture were significant, but not as strong as for other taxa. Moreover, it was noted that tamarack, a species for which the spatial distribution showed a concentrated pattern (Fig. 2f), had most of its bearing trees found on very poorly drained and organic soils. The degree of association between tamarack and soil conditions was not quantified because of its small sample size.

DISCUSSION

Approaches for presettlement vegetation reconstruction

Continuous representation of tree species

The spatially continuous representations of individual tree species that were created using the geostatistical method of indicator kriging showed spatial patterns that were not apparent in discrete point representations, or that were not revealed in summary tables of species frequencies typically used in presettlement vegetation reconstruction.

The distinct advantage of the indicator kriging approach is that it is able to employ fine-scale information about bearing trees because the semi-variogram for a species will be influenced if all of the bearing trees at a survey corner are of the same species, resulting in predictions of high probability occurrence for the species. This pattern, however, cannot be easily communicated using discrete points. In non-kriged point maps, bearing-tree locations are too near their survey corners to be clearly illustrated, and, hence, single or multiple bearing trees of the same species at a survey corner will often be seen as only one point. Therefore, the foci of abundance for the two species beech and sugar maple in the discrete point maps were not as apparent as those shown in the continuous representation from indicator kriging (Fig. 2a,b).

On the other hand, there are two potential points of concern regarding the use of geostatistics in presettlement vegetation reconstruction. First, just as for all the other interpolation methods, the accuracy of an interpolated surface from the geostatistical approach can be questioned in areas with sparse or no data, such as the centre of a township. This should not be a major issue that would alter the general conclusions made in this study because the private HLC surveys are a relatively systematic sampling of the landscape and the presettlement vegetation was reconstructed over a large area. In addition, an analysis based on the geostatistical approach employed herein suggests that the average prediction errors do not appear to exhibit any general pattern of increasing with the distance away from the township boundaries (Wang, 2004). It should be recognized, however, that uncertainties remain with regard to the predictions of species occurrence.

Second, the geostatistical method assumes spatial autocorrelation, which probably is not a fair assumption in complex terrain where abrupt changes in the underlying environmental factors may occur (Black *et al.*, 2002). Although the study area of western New York is large enough to have marked spatial variation in soils and physiography (Fig. 1a) that lead to variations in forest composition, the landscape is still quite flat compared with most parts of the north-eastern USA. Thus, the study area provides a situation in which forest composition might change quite gradually through space, allowing the use of geostatistical techniques that are suitable for describing spatial continuity (Burrough, 1996). An alternative approach would be to use a form of co-kriging that takes into account environmental variables. This approach, however, is not advisable because selecting appropriate environmental variables can be subjective, and using environmental factors to determine tree species distribution eliminates the independence of the tree species and site conditions, thereby inhibiting the investigation of vegetation–site relationships (Manies & Mladenoff, 2000; Wang, 2005).

Classification of vegetation associations

The spatially continuous representations of vegetation associations, based on classifications of overlaid probability maps of individual species, present a distinct advantage over previous approaches. As with the species maps, the gaps between data points are filled and the continuous distributions of vegetation associations (Fig. 3) are easier to interpret than discontinuous distributions of numbers or symbols that indicate different vegetation types. For example, the association of oak–white pine–chestnut can be easily recognized in the upland of the Allegany Reservation in the continuous representation (Fig. 3b). A similar vegetation type, black oak–white oak–chestnut–white pine, identified by Seischab (1992) is not, however, evident in the map of the 14 vegetation community types represented using different numbers along the survey lines. This study thus provides visually effective reconstructions compared with prior PLSR research.

The rule-based and statistically clustered methods employed in this study have the advantage of quantitatively summarizing vegetation compositional patterns in a single image, as opposed to manually and subjectively grouping species into associations and delineating the boundaries between the adjacent vegetation types. Subjective decisions are, however, involved in the two approaches and may influence the reconstructed vegetation patterns and diminish the objectivity of the results. The rule-based approach requires the determination of the cutoff value in advance, while the statistically clustered approach requires the assignment of some arbitrary number of groups that will be created. In general, the patterns reconstructed from the two approaches show strong similarity, as they were based on the same kriging procedure. Some of the differences between the rule-based and statistically clustered reconstructions might arise from the use of normalized data for cluster analysis. If non-normalized data had been used, fewer clusters would have appeared since most of the cells had high probability predictions for beech. Conversely, if normalized data had been used for the rule-based approach, this might have produced more combinations of different vegetation classes than the results from non-normalized data, which already generated more than 200 classes with the cutoff of 0.3.

Despite all these issues, the use of rule-based and statistically clustered approaches allows the reconstruction of presettlement vegetation using various classification schemes. The rule-based approach would be useful for examining local variants or transitional patterns because clustering may produce groupings based on species co-occurrence over the whole study area and hence mask local variations (Batek *et al.*, 1999). However, the clustering approach would be useful for comparing vegetation distributions from two periods of time. Common clusters can then be obtained by clustering the data sets of two different time periods as a group, and vegetation change can then be studied by examining the distribution change of these clusters.

Patterns of the presettlement vegetation in western New York

Effectiveness of the survey records

The presettlement vegetation in western New York was reconstructed using the coarsely resolved township data surveyed over a relatively short 3-year time period in which little vegetation change is likely to have occurred, and during a time period when the land was thinly populated, with little European settlement.

The tally of 8792 bearing trees suggested that beech and sugar maple were dominant in the presettlement forest composition of western New York. There have been concerns about surveyor bias towards the selection of bearing trees that were large, long-lived, or had thin and highly visible bark (Bourdo, 1956). From this perspective, the overwhelming dominance of beech, a species with thin, smooth bark that could be easily inscribed by surveyors, might indicate such a bias, and result in the bearing-tree data base being inaccurate in its representation of the relative abundance of trees. However, other reconstructions have also indicated a high presettlement proportion of beech in some areas of the north-eastern USA (Siccama, 1971; Seischab & Orwig, 1991; Cogbill *et al.*, 2002). At a large spatial extent of several states, the reconstructed presettlement vegetation suggests the tremendous dominance of beech over all northern New England (Cogbill *et al.*, 2002). At an intermediate extent of part of a state, the reconstruction of the western New York study area using the PLSR line descriptions (i.e. Seischab, 1992), which are considered free from the potential biases associated with selecting bearing trees (Almendinger, 1997), indicates that the five most frequent taxa in the HLC land were, in descending order, beech, sugar maple, hemlock, basswood, and elm, the same rank ordering as obtained in this study of bearing trees (Table 1). At a small extent of a county within the study area, the 'History of the City of Buffalo and Erie County' in 1620 noted that the county was covered with beech and maple in the valleys (Smith, 1884). Similarly, the presettlement vegetation reconstructed by Gordon (1940) for Cattaraugus County in south-western New York also noted the abundance of beech in the finely resolved lot-level data. The results from different spatial extents are thus in agreement with the reconstruction of this study that the presettlement vegetation of western New York was dominated by beech and sugar maple. They further support the contention made in prior studies that bearing-tree data are a statistical sample and the relative frequencies are a consistent and unbiased estimate of the overall composition of the presettlement vegetation (Almendinger, 1997; Cogbill, 2000).

Ecological interpretation

This research enhances the understanding of vegetation–site relationships in the presettlement forests of western New York. A visual comparison of species distribution and the origin of soils in prior research (Seischab, 1992) suggested that some of the dominant species, such as beech, sugar maple, and basswood, occurred on almost all soil types. The investigation of this study quantifies the degrees of association between species and soil conditions, suggesting that soil texture and drainage were both important site conditions in determining the distribution of individual species. Although beech and sugar maple were widely distributed, they did demonstrate different site preferences (Fig. 4). Compared with beech and sugar maple, basswood was positively associated with a wider range of soil drainage and texture classes. In addition, the standardized residual values (Fig. 4) suggested that the response of various taxa to drainage, also an important site determinant for the presettlement vegetation in other regions (Whitney & Steiger, 1985), was more apparent than those to texture. The research thus supports the findings in prior studies that the environmental factors of climate and soil played the major role in determining the presettlement vegetation patterns of the north-eastern USA (Russell, 1983; Cogbill *et al.*, 2002).

This study not only provides more visually effective reconstructions, but also shows vegetation associations at a much finer spatial resolution than did prior research (Seischab, 1992), offering more insights into the spatial pattern of presettlement vegetation. For example, a prominent feature in the Erie–Ontario Lowland was the east–west-trending aggregates of vegetation associations (Fig. 3), probably related to the pattern of surficial geology (*cf.* Cadwell, 1988) and that of soil drainage analysed in this study. In the spatial pattern reconstructed using the rule-based approach (Fig. 3a), the vegetation changed gradually from the aggregates of black ash, elm, and tamarack-related associations in the north, into a transitional area mixed with unnamed minor associations and areas that did not belong to any association type, before it changed into the beech- and sugar maple-dominated forests in the south. The existence of this transition between the associations of black ash, elm, and tamarack in the north and the large beech- and sugar maple-dominated cluster in the south suggested an ecotone of the region. The spatial pattern reconstructed using the clustering approach (Fig. 3b) showed a similar pattern. Several associations, such as the black ash–elm and the hemlock–northern hardwood associations, occurred more frequently on one side or the other of this transitional area.

Historical baseline

Based on the reconstruction in this study, the degree to which forest composition has changed over the last 200 years can be assessed using contemporary USDA Forest Service Forest Inventory and Analysis (FIA) data. Compared with the findings in Y.-C. Wang *et al.* (unpublished data), it is apparent that the relative frequencies of the three most abundant species, beech, sugar maple, and hemlock, have all declined significantly over the past two centuries. In addition, disease has devastated American chestnut, and probably contributed to the decline of beech, which is much less common today (4.6%) than it was in the presettlement forest (37.0%). Another possible cause for the decline in beech since 1800 can be attributed to its poorer dispersal ability than that of most associated hardwood species when forest stands are heavily cut (Dyer, 2001). It has been noted that, by the 1920s, most of the presettlement forest in the southern part of the study area had been logged (Seischab, 1993). Heavy logging results in fewer beech in the new stand than in the old, and repeated clear-cutting on short rotations may nearly eliminate the species (USDA Forest Service, 1965).

Sugar maple is now the most dominant species in the region, but its relative frequency has also declined, from 21.0% to 15.3%. The dominance of this species in the current forest composition and the decline of it since the presettlement might both be related to the dieback of beech. On the one hand, the immediate result of widespread death of beech would be an increase in sugar maple importance in the beech–maple association, the most common vegetation association in the presettlement (Table 3, Fig. 3b). On the other hand, the forest gaps created by the death of beech would have an influence on sugar maple. It has been observed in the Allegany Reservation of the study area that numerous sugar maples exhibited the effects of sun scalding (Seischab, 1993).

The dominant taxa in today's forest also include red maple, a shade-tolerant species that has increased in relative frequency from 2.6% to 10.3%. Other taxa that have increased significantly are shade-intolerant, early successional species, including ash from 6.0% to 13.2%, poplar (*Populus* spp.) from 0.5% to 8.9%, and black cherry from 0.5% to 6.7%, a change similar to that observed in south-eastern Ohio (Dyer, 2001). In summary, the species composition of the forests of western New York has changed significantly since the time of European settlement, reflecting a regime of increased forest disturbance and the effects of introduced pathogens. The present forests in western New York appear to be more diverse in their species composition than those of 200 years ago. The 10 most dominant species constituted 90% of the trees recorded in the presettlement survey, as opposed to 75% recorded in the current FIA data (Y.-C. Wang *et al.*, unpublished data).

ACKNOWLEDGEMENTS

I wish to thank Chris Larsen for valuable discussions, and Chen-Chieh Feng and Canice Chua for their technical assistance. Constructive comments on earlier drafts of this manuscript were provided by Ling Bian, Mark Cowell, David Mark, Dov Sax, and two anonymous referees. This research was supported by the Mark Diamond Research Fund from SUNY-Buffalo and in part by the Academic Research Fund from the National University of Singapore.

REFERENCES

Almendinger, J. (1997) *Minnesota's bearing tree database. Biological Report, no. 56*. Minnesota Department of Natural Resources, St Paul, MN.

Austin, M.P. & Smith, T.M. (1989) A new model for the continuum concept. *Vegetatio*, **83,** 35–47.

Barrett, L.R., Liebens, J., Brown, D.G., Schaetzl, R.J., Zuwerink, P., Cate, T.W. & Nolan, D.S. (1995) Relationships between soils and presettlement forests in Baraga County, Michigan. *American Midland Naturalist*, **134,** 264–285.

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

Black, B.A., Foster, H.T. & Abrams, M.D. (2002) Combining environmentally dependent and independent analyses of witness tree data in east-central Alabama. *Canadian Journal of Forest Research*, **32,** 2060–2075.

Bourdo, E.A. (1956) A review of the general land office survey and of its use in quantitative studies of former forest. *Ecology*, **37,** 754–768.

Braun, E.L. (1950) *Deciduous forests of eastern North America*. Hafner Publishing, New York, NY.

Broughton, J.G. (1966) *Geology of New York*. University of the State of New York, State Education Department, New York State Museum and Science Service, Albany, NY.

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 Science*, **12,** 105–129.

Bryer, M., Maybury, K., Adams, J.S. & Grossman, D. (2000) More than the sum of the parts: diversity and status of ecological systems. *Precious heritage: the status of biodiversity in the United States* (ed. by B.A. Stein, L.S. Kutner and J.S. Adams), pp. 201–238. Oxford University Press, Arlington, VA.

Burrough, P.A. (1996) Natural objects with indeterminate boundaries. *Geographic objects with indeterminate boundaries* (ed. by P.A. Burrough and A.U. Frank), pp. 3–28. Taylor & Francis, London.

Cadwell, D.H. (1988) *Surficial geologic map of New York, Niagara sheet. New York State Geological Survey Map and Chart Series, no. 40*. University of the State of New York, The State Education Department, Albany, NY.

Campbell, J. (2001) *Map use and analysis*, 4th edn. McGraw-Hill, New York, NY.

Cogbill, C.V. (2000) Vegetation of the presettlement forests of northern New England and New York. *Rhodora*, **102,** 250–276.

Cogbill, C.V., Burk, J. & Motzkin, G. (2002) The forests of presettlement New England, USA: spatial and compositional patterns based on town proprietor surveys. *Journal of Biogeography*, **29,** 1279–1304.

Cox, C.B. & Moore, P.D. (1993) *Biogeography: an ecological and evolutionary approach*, 5th edn. Blackwell Science, Oxford.

Curtis, J.T. (1959) *The vegetation of Wisconsin*. University of Wisconsin Press, Madison, WI.

Dale, V.H. & Haeuber, R.A. (2001) *Applying ecological principles to land management*. Springer-Verlag, New York, NY.

Delcourt, H.R. & Delcourt, P.A. (1974) Primeval magnolia–holly–beech climax in Louisiana. *Ecology*, **55,** 638–644.

Dyer, J. (2001) Using witness trees to assess forest change in southeastern Ohio. *Canadian Journal of Forest Research*, **31,** 1708–1718.

Easterling, D.R., Karl, T.R., Mason, E.H., Hughes, P.Y. & Bowman, D.P. (1996) *United States historical climatology network (US HCN) Monthly temperature and precipitation data. ORNL/CDIAC-87, NDP-019/R3*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, TN.

Edinger, G.J., Evans, D.J., Gebauer, S., Howard, T.G., Hunt, D.M. & Olivero, A.M. (2002) *Ecological communities of New York State*, 2nd edn. A revised and expanded edition of Carol Reschke's Ecological Communities of New York State (Draft for review). New York Natural Heritage Program, New York State Department of Environmental Conservation, Albany, NY.

Egan, D. & Howell, E.A. (2001) *The historical ecology handbook: a restorationist's guide to reference ecosystems*. Island Press, Covelo, CA.

Forman, R.T.T. (1995) *Land mosaics: the ecology of landscapes and regions*. Cambridge University Press, New York, NY.

Gleason, H.A. (1926) The individualistic concept of plant association. *Bulletin of the Torrey Botanical Club*, **53,** 7–26.

Goodchild, M.F. (1992) Geographical data modeling. *Computers and Geosciences*, **18,** 401–408.

Goodchild, M.F. (1994) Integrating GIS and remote sensing for vegetation analysis and modeling: methodological issues. *Journal of Vegetation Science*, **5,** 615–626.

Gordon, R.B. (1940) The primeval forest types of southwestern New York. *New York State Museum Bulletin*, **321,** 3–102.

Grimm, E.C. (1984) Fire and other factors controlling the big woods vegetation of Minnesota in the mid-nineteenth century. *Ecological Monographs*, **54,** 291–311.

Haberman, S.J. (1973) The analysis of residuals in cross-classified tables. *Biometrics*, **29,** 205–220.

He, H.S., Mladenoff, D.J., Sickley, T.A. & Guntenspergen, G.G. (2000) GIS interpolations of witness tree records (1839–1866) for northern Wisconsin at multiple scales. *Journal of Biogeography*, **27,** 1031–1042.

Howell, D.L. & Kucera, C.L. (1956) Composition of pre-settlement forests in three counties of Missouri. *Bulletin of the Torrey Botanical Club*, **83,** 207–217.

Johnston, C.A. (1998) *Geographic information systems in ecology*. Blackwell Science, Malden, MA.

Kemp, K.K. (1997) Fields as a framework for integrating GIS and environmental process models. Part 1: Representing spatial continuity. *Transactions in GIS*, **1,** 219–233.

Küchler, A.W. (1964) *Potential natural vegetation of the conterminous United States. Special Publication no. 36*. American Geographical Society, New York, NY.

Küchler, A.W. & Zonneveld, I.S. (1988) *Vegetation mapping*. Kluwer Academic Publishers, Dordrecht.

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

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

MacDonald, G.M. (2003) *Biogeography: space, time, and life*. John Wiley & Sons, New York.

Manies, K.L. & Mladenoff, D.J. (2000) Testing methods to produce landscape-scale presettlement vegetation maps from the U.S. public land survey records. *Landscape Ecology*, **15,** 741–754.

McIntosh, R.P. (1962) The forest cover of the Catskill Mountain region, New York, as indicated by land survey records. *American Midland Naturalist*, **68,** 409–423.

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

Russell, E.W.B. (1983) Indian-set fires in the forests of the northeastern United States. *Ecology*, **64,** 78–88.

Schulte, L.A. & Mladenoff, D.J. (2005) Severe wind and fire regimes in northern Wisconsin (USA) forests: historical variability at the regional scale. *Ecology*, **86,** 431–445.

Schulte, L.A., Mladenoff, D.J. & Nordheim, E.V. (2002) Quantitative classification of a historic northern Wisconsin (USA) landscape: mapping forests at regional scales. *Canadian Journal of Forest Research*, **32,** 1616–1638.

Seischab, F.K. (1992) Forests of the Holland Land Company in western New York, circa 1798. *New York State Museum Bulletin*, **484,** 36–53.

Seischab, F.K. (1993) *Forest of the Allegany State Park Region: 200 years of forest change*. Unpublished report to the Allegany State Park, New York, USA.

Seischab, F.K. & Orwig, D. (1991) Catastrophic disturbances in the presettlement forests of western New York. *Bulletin of the Torrey Botanical Club*, **118,** 117–122.

Shanks, R.E. (1953) Forest composition and species association in the beech–maple forest region of western Ohio. *Ecology*, **34,** 455–466.

Siccama, T.G. (1971) Presettlement and present forest vegetation in northern Vermont with special reference to Chittenden County. *American Midland Naturalist*, **85,** 153–172.

Smith, H.P. (1884) *History of the city of Buffalo and Erie County: with illustrations and biographical sketches of some of its prominent men and pioneers*. D. Mason & Co., Syracuse, NY.

Sokal, R.R. & Rohlf, F.J. (1995) *Biometry: the principles and practice of statistics in biological research*. Freeman, New York, NY.

Spurr, S.H. (1951) George Washington, surveyor and ecological observer. *Ecology*, **32,** 544–549.

Strahler, A.H. (1978) Binary discriminant analysis: a new method for investigating species–environment relationship. *Ecology*, **59,** 108–116.

USDA Forest Service (1965) *Silvics of forest trees of the United States (compiled and revised by H.A. Fowells)*. US Department of Agriculture, Agriculture Handbook no. 271, Washington, DC.

USDA Natural Resources Conservation Service (USDA NRCS) (1995) *State Soil Geographic (STATSGO) data base: data use information*. US Department of Agriculture, Natural Resources Conservation Service, Miscellaneous Publication no. 1492, Washington, DC.

Veatch, J.T. (1925) Soil maps as a basis for mapping original forest cover. *Michigan Academic Science*, **15,** 267–273.

Wang, Y.-C. (2004) A geographic information approach to analyzing and visualizing presettlement land survey records for vegetation reconstruction. PhD thesis, University at Buffalo, The State University of New York.

Wang, Y.-C. (2005) Presettlement land survey records of vegetation: geographic characteristics, quality and modes of analysis. *Progress in Physical Geography*, **29,** 568–598.

Wang, Y.-C. & Larsen, C.P.S. (2006) Do coarse resolution U.S. presettlement land survey records adequately represent the spatial pattern of individual tree species? *Landscape Ecology*, **21,** 1003–1017.

Whitney, G.G. (1982) Vegetation–site relationships in the presettlement forests of northeastern Ohio. *Botanical Gazette*, **143,** 225–237.

Whitney, G.G. & DeCant, J. (2001) Government land office survey and other early land surveys. *The historical ecology handbook: a restorationist's guide to reference ecosystems* (ed. by D. Egan and E.A. Howell), pp. 147–172. Island Press, Covelo, CA.

Whitney, G.G. & Steiger, J.R. (1985) Site-factor determinants of the presettlement prairie–forest border areas of north-central Ohio. *Botanical Gazette*, **146,** 421–430.

Whittaker, R.H. (1951) A criticism of the plant association and climatic climax concepts. *Northwest Science*, **25,** 17–31.

Whittaker, R.H. (1970) *Communities and ecosystems*. MacMillan, New York, NY.

Wuenscher, J.E. & Valiunas, A.J. (1967) Presettlement forest composition of the River Hills region of Missouri. *American Midland Naturalist*, **78,** 487–495.

Wyckoff, W. (1988) *The developer's frontier: the making of the western New York landscape*. Yale University Press, New Haven, CT.

BIOSKETCH

Yi-Chen Wang is an Assistant Professor in geography at the National University of Singapore. Her research interests include reconstruction of historical forest landscapes from land survey records; investigating patterns and processes of vegetation change; and applications of GIS and remote sensing technology in mapping and modelling land-cover change across multiple scales of space and time.

Editor: Dov Sax