The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

Title:

Presettlement Land Survey Records of Vegetation: Geographic Characteristics, Quality, and Modes of Analysis

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598. Please contact me at geowyc@nus.edu.sg for reprint.

Author name and address:

Yi-Chen WANG

Department of Geography

National University of Singapore

1 Arts Link

Singapore 117570

Email: geowyc@nus.edu.sg

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

Abstract:

Presettlement land survey records (PLSRs) of North America provide insights into the landscape before European disturbance and have been used in various ecological studies. The usefulness of the data, however, varies with their characteristics and qualities. The purpose of this paper is to examine the geographic characteristics of PLSRs, to review PLSR studies in the last five decades and trends in these studies based on the geographic characteristics that have been analyzed, and to investigate data quality issues concerning the PLSRs. The framework used in this examination is that of Geographic Information Science (GIScience), including geographic characteristics of space, theme, and time, and data quality components of lineage, positional accuracy, attribute accuracy, logical consistency, and completeness. The results show that different types of PLSRs have distinct geographic characteristics, especially regarding the spatial characteristics of shape and size/resolution. Prior studies have used PLSRs in six modes of analysis, with presettlement vegetation patterns and compositions analyzed most often. Data quality investigations suggest that whether the potential quality issues will influence an analysis depends on the study purpose and the spatial extent of interest. In cases of studying vegetation dynamics of small areas, the positional accuracy of landscape features is essential, while this issue may not impede large area reconstruction. Finally, thoughts concerning future research avenues are presented. This study uses a GIScience framework to provide a more systematic and comprehensive examination of the usefulness and limitations of PLSRs than previous studies, and can assist future research to employ the data most appropriately.

Keywords:

presettlement vegetation, geographic characteristic, data quality, General Land Office survey, Holland Land Company survey, metes and bounds survey

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

A I. Introduction

The majority of the landscape in the developed world has been disturbed by human occupation and use of the land (Brown, 1998a). To understand these impacts, knowledge of the original vegetation is desirable. Even in areas where old-growth forests are present, their current structure and floristic composition do not reflect their presettlement characteristics (Fensham and Fairfax, 1997) because old-growth forests have developed under protection while presettlement forests developed under a complex of climate, topography, and natural disturbances (Fralish et al., 1991).

Various sources of information, including pollen records, travelers' and settlers' accounts, historical photographs, and presettlement land survey records (PLSRs), have been employed in original vegetation reconstruction. Among them, the PLSRs are considered the most important in North America (Whitney, 1996). The PLSRs were collected between the late 17th and the early 20th centuries for land sale and settlement. Although the data were not collected for ecological purposes, they provide a representation of the vegetation that existed prior to significant European settlement.

The PLSRs can be categorized according to whether the shape of the survey system was irregular or regular; the regular survey systems can further be classified according to whether the survey agency was a private land company or public government. The PLSRs therefore have three types: (1) irregular metes and bounds surveys, (2) regular private land surveys, and (3) regular public land surveys. These data have enjoyed new use since the early 20th century as a source of data for presettlement vegetation studies, including presettlement plant community composition (Lutz, 1930; McIntosh, 1962; Nelson, 1997); patterns of natural disturbance (Lorimer, 1977; Grimm, 1984); vegetation-site relationships (Cowell, 1995; Abrams and McCay, 1996); vegetation dynamics (Siccama, 1971; White and Mladenoff, 1994); and human impacts since European settlement (Foster, 1992). The results from these studies have in turn been used to guide restoration of disturbed habitats (Galatowitsch, 1990; Radeloff et al., 1998). The number of theses and papers that use PLSRs in presettlement vegetation studies has grown dramatically in the past half century. It now totals several hundred references (Whitney and DeCant, 2001).

The usefulness of PLSRs for vegetation studies is limited, however, because the data were not collected for ecological purposes, and therefore do not have optimal data characteristics or qualities. Examinations of the limitations of the data were first conducted by Bourdo (1956) with respect to surveyors' procedures and biases. This critical study is the one most often cited in PLSR research. Throughout the 1990s, several other articles as to the usefulness and limitations of PLSRs were published. Applications and scientific uses of the data were summarized (Schulte and Mladenoff, 2001; Whitney and DeCant, 2001). Limitations and qualities of the data were discussed and evaluated, such as fraudulent surveys and misidentification of tree species (Galatowitsch, 1990; Whitney, 1996; Almendinger, 1997), the accuracy of forest composition representation (Bürgi and Russell, 2000), and the extent of variability among surveyors (Manies et al., 2001). A review on this topic, however, is still needed for three reasons. First, the usefulness and quality of PLSRs have been examined in an *ad hoc* manner as various researchers have analyzed and worked with the data. Previous studies have also focused mainly on issues related to the public land survey records rather than to the metes and bounds and the private land survey records. A more comprehensive and systematic assessment is required for the most appropriate uses of all three types of PLSRs. Second, although studies mentioning the three types of PLSRs have provided a summary of the uses of the data (Whitney and DeCant, 2001), no articles have tracked how researchers have used the PLSRs over the years. A review is necessary to investigate how researchers have made progress with advances of new theories, methods, and techniques in geography and ecology. Third, all the reviews have been published by ecologists outside of the geographic literature. Only the use of public land survey records has been briefly discussed in *The Professional Geographer* (Pattison, 1956). A review from a geographic perspective is desired because the PLSRs consist of wholly geographic information, and have been noted and used by geographers since the early 20th century (Trewartha, 1932).

The goal of this paper is to provide a systematic review of the geographic characteristics, applications, and limitations of PLSRs. The theoretical framework that I use is that of Geographic Information Science (GIScience). This framework has been developed over the past several decades to study geographic information (Sinton, 1978; Maguire and Dangermond, 1991; Chrisman, 1997). Since PLSRs consist wholly of geographic information, GIScience provides a most appropriate theoretical framework on which the geographic characteristics can be described, the applications can be reviewed, and the data qualities can be examined.

This paper contains five parts. First, a brief background on the presettlement land surveys will be provided. Second, the geographic characteristics of PLSRs will be elucidated. Third, the PLSR studies will be reviewed based on how they utilize the geographic characteristics. The evolution of each type of utilization will also be documented. Fourth, the data quality components used in GIScience will be used to investigate the qualities of PLSRs. Further use of the data and prompting trends in PLSR research will be identified at the end. This systematic analysis should allow future studies to employ PLSRs most thoroughly and appropriately.

A II. Background on the Presettlement Land Surveys

Presettlement Land Surveys were conducted to inventory land quality and to establish landmarks for land sale and settlement (Whitney, 1996). In the colonial period, irregular metes and bounds surveys were commonly used, especially towards the end of the 17th century. 'Metes' implies an act of metering, measuring, or assigning by measures; 'bounds' refers to property boundaries or the limiting extent of an ownership (Avery, 1967). The irregular nature of metes and bounds surveys was the result of the fact that landscape features such as rivers and ridges were often used as boundaries (Figure 1), and because early warrant holders attempted to obtain the best land (Whitney, 1996). The irregular shapes caused many of the resulting patents and deeds to have land gaps and overlaps, and this led to many boundary disputes (White, 1984).

Figure 1 here#

The boundary problems caused by the haphazard nature of the metes and bounds surveys prompted the advocacy of a rectangular survey system prior to land sale or settlement in unsettled regions. Traces of early rectangular units have been found in the New England town system in the 17th century (White, 1984) and in small areas within the larger areas surveyed using the metes and bounds system (Truesdell, 1908). The rectangular system was far superior to the irregular metes and bounds because it surveyed all of the land with no gaps (Whitney, 1996).

The Land Ordinance of 1785 was the legal origin of the rectangular system of public land surveys and the first legislation on the subject (Stewart, 1935). The Land Ordinance influenced successive legislations and survey instructions for the public land surveys from Ohio to California in subsequent decades (Wyckoff, 1988). The General Land Office (GLO), established in 1812, carried out most of the public land surveys. Land was typically divided into 9.6 × 9.6 km (6 × 6 mi) townships composed of 36 1.6 × 1.6 km (1 × 1 mi) sections (Figure 2).

Figure 2 here#

Private developers (e.g., the Ohio Land Company in southeastern Ohio and the Holland Land Company in western New York) were also aware of the advantage of the rectangular survey system (Wyckoff, 1988). A 9.6 × 9.6 km township system was often employed (Figure 3), though it was subdivided into various sized lots. The private land surveys were conducted in central and western New York, northwestern Pennsylvania, and northeastern and southeastern Ohio. Although they did not occupy a large extent as the public land surveys did, the private land surveys existed in a transition era between the irregular metes and bounds and the regular public land surveys, and therefore have unique characteristics from the other two survey systems.

Figure 3 here#

The PLSRs can therefore be categorized into three types: (1) the irregular metes and bounds, which represent the survey system in the colonial period, (2) the regular GLO surveys, which represent the typical public land surveys, and (3) the regular private land surveys, of which the Holland Land Company (HLC) survey of western New York is used as an example. The geographic characteristics of the three types of PLSRs are examined in the following section.

A

III. Geographic Characteristics of the Presettlement Land Survey Records

A GIScience framework is employed to review the geographic characteristics of PLSRs because the data have an inherently geographic nature. GIScientists recognize three main dimensions of geographic information: space, theme, and time (Sinton, 1978; Marble et al., 1984). Spatial information allows objects or phenomena to be located, thematic information provides aspatial characteristics about them, and temporal information indicates the time and age of each observation. The specific geographic characteristics that I have used to examine PLSRs are derived and synthesized from a review of GIScience references (Sinton, 1978; Maguire and Dangermond, 1991; Laurini and Thompson, 1992; Bonham-Carter, 1994; Martin, 1996; Chrisman, 1997; Schneider, 1997). Each characteristic is broken into sub-characteristics (Table 1). The list in Table 1 is not intended to exhaust all possibilities; only those most meaningful to the use of PLSRs are considered.

Table 1 here#

The three types of PLSRs are examined in terms of the geographic characteristics in Table 1. A detailed understanding of these characteristics first allows PLSR researchers to appreciate the different types of data available to them, and second, provides a complete framework by which the PLSR studies can be reviewed and the data quality issues can be addressed.

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

B 1. Space

Spatial characteristics include location, geometry, and topology, and are the key features that make geographic data differ from other types of data (Maguire and Dangermond, 1991).

C a. Location

Location refers to where an existing object is located. Two types of locational information in PLSRs are of particular interest: location of surveys and location of surveyed objects. Regarding the first type of information, it is important to note where the three types of surveys were located. The metes and bounds system was found in most of the original thirteen colonies, especially those in the American South (Avery, 1967). The private land surveys, such as the HLC survey in western New York, and others in central New York, northern Pennsylvania, and southeastern and northeastern Ohio (Wyckoff, 1988), were located immediately to the west of the metes and bounds surveys. The public land surveys were conducted further to the west, from Ohio to California (Pattison, 1964), including west of the Mississippi River and north of the Ohio River, plus Alabama, Mississippi, and portions of Florida (Avery, 1967).

The second type of locational information deals with aspects of where the surveyed objects such as survey posts, trees, and other landscape features were located, with trees being of particular interest for presettlement vegetation studies. The GLO surveyors erected posts at the intersection of section lines (section corners), the midpoint between section corners (quarter corners), and where section lines crossed navigable rivers or lakes (meander corners). At these survey corners, two to four nearby trees were blazed as witness trees. Witness trees to which the bearing and distance from the post were noted were known as bearing trees (Whitney and DeCant, 2001). Surveyors also recorded trees that fell along the survey lines as line trees. In the private HLC surveys, two to four bearing trees were mainly recorded at the survey corners along the township boundaries. In the metes and bounds surveys, significant trees along the survey lines were recorded as reference points. A post was occasionally used to mark a survey corner, but none or only one bearing tree was provided (McIntosh, 1962).

The locations of bearing trees are essential for mapping presettlement vegetation. They can be plotted directly on maps to show the presence or absence of individual species, and then grouped subjectively based on apparently homogeneous patterns of abundant species to show vegetation associations (Siccama, 1971; Barrett et al., 1995). In addition, they can be tied to specific soil associations or physiographic positions to reestablish vegetation-site relationships in presettlement forests (Cowell, 1995).

C b. Geometry

Two important geometric characteristics of PLSRs are shape and size/resolution. Shape is considered at the level of the outline of a land parcel and the features located within it. The shape of the land parcel was often irregular in the metes and bounds, but more regular in the GLO (often square) and the private land (rectangular or square) surveys. The shape of a spatial object also describes an abstraction of its geometric structure as a point, line, or polygon (Schneider, 1997). The shape of a tree can thus be modeled as a point and the shape of a river as a line, although in reality they both are three-dimensional.

PLSRs contain the geometric features of points, lines, and polygons. Point features include survey corners/posts, bearing/witness trees, and other points along the survey lines. Line features are line descriptions, the narrative descriptions of vegetation, soil quality, and occurrence of disturbance along the survey lines. Polygon features include landscape features of lakes, swamps, timber, prairies, and burnt areas drawn on township plats by the surveyors. The metes and bounds often include only point features of survey posts and bearing trees. The GLO and the private HLC surveys have all three geometric features but to different degrees. The GLO surveys provide line descriptions at 0.8 or 1.6 km (0.5 or 1 mi) intervals (Whitney, 1986), while the private HLC surveys provide the information at several segments of each surveyed mile (Seischab, 1992). For polygon features, detailed township plats showing the distribution of landscape features accompanied the GLO surveys (Whitney, 1996), while the private HLC surveys only included sketch maps describing terrain features crossed survey boundaries (Wyckoff, 1988). It is important to know the geometric features and their availability in PLSRs because the lack of certain geometric features might constrain the potential utilization of the data.

Size consists of measurements such as the length of a line, or the perimeter and the area of a region. Resolution, or grain in landscape ecology, is the size of the smallest addressable spatial unit (Laurini and Thompson, 1992) such as pixels in remote sensing or sampling intervals in ecology (Lam and Quattrochi, 1992). Spatial resolutions may greatly influence the detection and characterization of presettlement landscape patterns.

Size and resolution vary greatly among the three PLSRs. The rectangular systems of the GLO and the private HLC surveys produced more equal-sized polygons compared to the variety of land sizes of the metes and bounds surveys. The GLO land was divided into 9.6 × 9.6 km townships, 1.6 × 1.6 km sections (640 acres), and 0.8 × 0.8 km quarter sections (160 acres) (Figure 2). Not until 1832 was a standard 0.16 km^2 (40 acre) parcel of public land made available (Wyckoff, 1988). The private HLC land was divided into different rectangular townships with 9.6 × 9.6 km mostly used, and then subdivided into sections or lots ranging from 0.16 to 2.59 km^2 (40 to 640 acres). The various smaller lots of private developers reflected the purpose of attracting small retail purchasers (Wyckoff, 1988).

The sampling intervals between bearing trees were multi-resolutioned: finely resolved for bearing trees around a survey corner, and coarsely resolved between survey corners. Overall, the sampling interval was coarse, thus resulting in a low-density of sampled trees for PLSRs. The estimated density of sampled trees for the GLO surveys in Upper Michigan was about 5 trees per km^2 (Brown, 1998a) and ranged from about 3 to 7 trees per km^2 in other regions of the United States (Schwartz, 1994; Delcourt and Delcourt, 1996). The HLC survey provided a smaller number of sampled trees per km^2 than the GLO survey because it was along the township boundaries only. The estimated density for the private HLC township survey was about 1.5 trees per km^2, but would be similar to the estimation for the GLO survey if the bearing trees for the HLC lot corners were included. The sampling intervals of the metes and bounds surveys were different from each other due to the haphazard nature of the survey system. The general density is therefore more difficult to calculate than the GLO and the private surveys, but averaged 1.5 to

2 trees per km^2 in the Catskill Mountain region of New York (McIntosh, 1962) and 3 trees per km^2 in eastern West Virginia (Abrams and McCay, 1996).

The coarse resolution of PLSRs makes extrapolation of vegetation associations problematic except in cases of extremely homogeneous cover (Cowell, 1995). Studies, however, have suggested that the tree samples in PLSRs are best applied at a broad spatial extent such as the county or state level because at such spatial extents the fine-scale complexity of the natural vegetation boundaries diminishes (Manies and Mladenoff, 2000).

C c. Topology

The topological characteristics of PLSRs include adjacency and connectivity. Adjacency represents the spatial relationship of neighboring objects (Laurini and Thompson, 1992). If bearing trees were selected based on their adjacency to the survey posts (i.e., they were the closest trees to the survey posts), then the distances between the trees and the posts can be used to obtain forest density (Cottam and Curtis, 1956), one of the primary measures used to characterize vegetation, and can further be combined with measures of spatial distribution to infer habitat preference and competitive dynamics (Barbour et al., 1999).

Indeed, in most PLSRs, trees adjacent to the posts were probably selected as bearing trees. The 1855 GLO surveyor's instruction documented, "courses (from the post) must be taken and distances measured to two or more adjacent trees in opposite directions, as nearly as may be, and these are called bearing trees." (Bourdo, 1956: 757). However, exceptions can still be found.

Some GLO surveyors were asked to choose bearing trees according to tree size and longevity, regardless of the adjacency (Grimm, 1984). Another adjacency characteristic found in the GLO and the private HLC surveys is the distribution of land parcels that were regularly tiled based on the reference axes of the principal meridian (a true north-south line) and base line (a true east-west line that corresponds to a parallel of latitude) (Avery, 1967). Townships are therefore adjacent to other townships, and sections are adjacent to other sections.

Connectivity provides information on how points are related using the operators of 'from-' and 'to-.' This is different from the use of the term connectivity in landscape ecology, where it signifies the degree to which all nodes in a system are linked by corridors (Forman and Godron, 1986). The processes by which surveyors surveyed the land demonstrate the connectivity characteristic. The surveyors started at a survey corner recording nearby bearing trees, and then from that corner, moved on to another corner recording nearby bearing trees. This characteristic is essential because it enables researchers to derive absolute coordinates of bearing trees that were not directly recorded in PLSRs.

B 2. Theme

Thematic characteristics identify the spatial phenomena or objects apart from their locations. There are various ways in which thematic characteristics can be classified. Stevens' (1946) four levels of measurement have become a basis for social science methods and a framework for GIS (Maguire and Dangermond, 1991; Martin, 1996), and therefore are employed here. The four levels of measurement are nominal, ordinal, interval, and ratio. Since no PLSRs use the interval level of measurement, it is excluded from the following discussion. Additionally, Chrisman (1997) noted other measurements that do not fit into Stevens' (1946) scheme. Chrisman's 'cyclical measurement' is incorporated because it is applicable to the PLSRs.

C a. Nominal

At the nominal level of measurement, each category has a distinct nature. The nominal measurement in all three types of PLSRs was tree species. Additionally, the GLO and the private HLC surveys recorded natural landscape features such as rivers, swamps, and lakes, as well as man-made landscape features such as roads and villages. Information on this level of measurement provides a general picture of the presettlement environment, and shows whether or not the study area had experienced minimal human disturbance.

C b. Ordinal

Ordinal measurements involve observations that are ranked according to relative position on a scale, with unequal intervals between units (Bonham-Carter, 1994). Some ordinal categories use a semantic scale of words. For example, soils are ordered from 'poorly drained' through 'somewhat poorly drained' on to 'well-drained' and 'excessively drained' (Chrisman, 1997). Ordinal measurements were usually absent in the metes and bounds, but were found in the GLO and the private HLC surveys. Surveyors classified soils according to their drainage (e.g., poorly drained or well-drained) and quality for cultivation (i.e., first rate, second rate, third rate, and unfit for cultivation). Moreover, the GLO surveyors were required to name tree species in order of their prevalence in the line descriptions (Bourdo, 1956). The first mentioned species was thus supposed to be the most dominant along that line segment. This characteristic provides an independent means of verifying the values obtained in the bearing tree counts to test for biases in the bearing tree selection (Whitney, 1986).

C c. Ratio

Ratio measurements have an absolute zero point and a unit, such as surveyor's measures in chains and links or feet and inches. They were used to record distances from bearing trees to survey corners, lengths of the survey line segments in line descriptions, widths of streams and rivers, and estimated diameters of bearing trees. The GLO and the private HLC surveys provided most of these ratio measurements, but estimated diameters were only found in the GLO surveys.

Among these ratio measurements, distances between bearing trees and corners and estimated tree diameters are of particular interest because they provide insights into the presettlement vegetation structure. Distances can be used to distinguish between prairie, savanna, open forest, and closed forest ecosystems (Anderson and Anderson, 1975). Diameter class distributions (i.e., size-class structure) can be used as a predictive tool in both population and community ecology and have been proven useful in deciphering the history of forest stands, especially for species whose size correlates well with age (Whitney, 1996; Barbour et al., 1999). An inverse J-shaped distribution with many seedlings and saplings and fewer larger trees suggests a high probability of a population maintaining itself as part of a climax community (Barbour et al., 1999).

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

Deviation from the theoretically expected inverse J-shape distribution might be attributable to selective logging, the introduction of pathogens, and fire cessation (Whitney, 1996).

C d. Cyclical

Cyclical measurements are bounded within a range and repeat in some cyclical manner (Chrisman, 1997). Angles are an example; although they seem to be ratio, in the sense that there is a zero and an arbitrary unit, they return to their origin. A common cyclical measurement in all three types of PLSRs is the compass directions of bearing trees from the survey corners. This information is essential for researchers to trace the locations of bearing trees.

B 3. Time

Time is the third characteristics of geographic data. Its two sub-characteristics are time of observation and age of observed object (Bonham-Carter, 1994).

C a. Time of observation

Time of the observation refers to the time when the data were collected; in PLSRs, it thus refers to the time of the survey. Knowledge of this characteristic indicates for how long European settlement may have disturbed the landscape. The time of settlement varied for different regions, but the time of the surveys relative to settlement was generally constant. The metes and bounds surveys prevailed in the colonial period, around the 17^{th} and the early 18^{th} centuries. The private land surveys were conducted during the late 18th to the early 19th centuries. The public land surveys were conducted during the late 18th to the early 20th centuries. Furthermore, surveys for township division were conducted first, and section and lot subdivision were conducted at a later date. For example, the township boundaries of the Ohio Company Purchase in southeastern Ohio were surveyed from 1788 to 1789, but many of the interior lines were not surveyed until from 1796 to 1802 (Dyer, 2001). For a larger area such as a state, the overall vegetation was sampled over a shifting period spanning more than several decades. Although the PLSRs of a given area were not gathered within the same year, and in some areas were collected at a time when the land was occupied, and the vegetation altered, to some degree by Native Americans, these data still effectively represent a single ecological period of presettlement (Brown, 1998a; Schulte and Mladenoff, 2001).

C b. Age of observed object

The surveyors did not directly measure or suggest ages of observed objects, but they recorded age-related information that can be used to infer ages of objects. For example, the occurrence of burnt land or windfalls recorded in line descriptions of the GLO and the private land surveys can be used to obtain the disturbance return time, the average span of time between consecutive disturbances at a given site (Seischab and Orwig, 1991). Under certain situations, this is equivalent to the time it takes to disturb an area equal to the size of the region under consideration. We can infer the disturbance return time by assuming that the effect of disturbance would only be observed for a certain number of years after its occurrence (Canham and Loucks, 1984). For example, Canham and Loucks (1984) and Whitney (1986) suggested that large-scale windfalls were still recognizable by surveyors for up to 15 to 30 years after their occurrence; in Minnesota, fires were recognizable for at least 20 years in the pine forests (Whitney, 1986). These estimated numbers of years, combined with the length of the disturbance intercepting the survey lines as recorded in line descriptions, are then used to obtain the disturbance return time or frequency of the disturbance.

A IV. Modes of Geographically Analyzing Presettlement Land Survey Records

With the geographic characteristics of PLSRs now described, we can review previous studies according to the characteristic they tackle, i.e., analyzing spatial, thematic, or temporal characteristics (Sinton, 1978). In the case of mapping presettlement vegetation, the spatial characteristic of bearing tree locations is analyzed to create vegetation distribution on a map. The geographic characteristics of space, theme, and time, thus serve as a framework to systematically review studies using PLSRs, and to allow researchers to realize that PLSRs are open to many modes of analysis. Since each geographic characteristic contains multiple sub-characteristics, there are many variations on these major modes of analysis, and they will be discussed. Furthermore, the trend of each mode of analysis will be reviewed with respect to advances of new theories, methods, and techniques in geography and ecology.

B 1. Analysis of spatial characteristics

C a. Creation of presettlement vegetation maps

Maps of presettlement vegetation communicate the amount of geographic variation in vegetation conditions. Mapping presettlement vegetation conditions is a major use of PLSRs. The spatial characteristic of locations of bearing trees and other landscape features is analyzed to delineate presettlement species or community distribution on a map using point symbols or polygons with subjectively chosen boundaries. Community types have been identified using subjective analysis (Finley, 1976; Schafale and Harcombe, 1983) and quantitative methods such as two-way indicator species analysis (TWINSPAN) and detrended correspondence analysis (DCA) (Seischab, 1992). Studies of this type are further reviewed according to the spatial characteristics of study areas, the geometric features of PLSRs, and the representation of vegetation distribution.

The spatial characteristics of a study area include descriptive factors such as spatial extent (size), location, and physiography. For spatial extent, the PLSRs have been used to reconstruct areas as small as several sections (Liegel, 1982) or a few townships (Delcourt and Delcourt, 1974), to part of a State (Grimm, 1984; Seischab, 1992), a whole State (Finley, 1976), or several states (Whitney, 1996), but most studies tend to deal with a county (Gordon, 1940; Siccama, 1971; Anderson and Anderson, 1975; Schafale and Harcombe, 1983). This may be a trade-off between sample adequacy (smaller areas of a few sections reduce the number of trees used in analysis) and data process duration (larger area of a state needs time commitment and data sharing and standardization). The locations of the study areas are throughout the USA, and in part of the Canadian Province of Ontario (Jackson et al., 2000), but have primarily been in the Northeast and Midwest of the USA (Sear, 1925; Lutz, 1930; Gordon, 1940; McIntosh, 1962; Grimm, 1984; Loeb, 1987; Brown, 1998b; Dyer, 2001; Cogbill et al., 2002). The physiographic types that have been most studied using the PLSRs are flat-lands such as river valleys (Delcourt and Delcourt, 1977; Nelson et al., 1994; Schulte and Barnes, 1996). Steep areas are seldom examined due to the fact that they are less likely to be sampled by witness trees (McIntosh, 1962; Liegel, 1982).

Points are the geometric features used in most PLSR studies, although lines and polygons are also used. The reasons for this are threefold. First, the exact positions of point data are easier to be located. Second, line descriptions are not as amenable to a conversion to quantitative measures as are point observations. Third, polygon data, as represented on township plats, offer too much data to allow large areas to be time-effectively studied; they are mainly used at township level reconstructions (Whitney, 1987) although state level reconstructions have been achieved (Marschner, 1974; Finley, 1976). Despite the many advantages of the point data, two disadvantages were noted. First, the coarse resolution of bearing trees impedes the reconstruction of small areas (less than a few townships) or the representation of complex landscapes (Manies and Mladenoff, 2000). This problem is greater when using the point data from the metes and bounds and private land surveys since these surveys include fewer bearing trees per km^2 than the GLO surveys. To minimize this problem, some studies have incorporated other historical records or photos (Siccama, 1971; Russell, 1981). Second, unlike the polygon features which offer rich information on landscape features, point data mainly offer information on trees. Therefore, the choice of geometric features for mapping presettlement vegetation will depend on the spatial extent and purpose of the study in addition to the availability of the geometric features of PLSRs.

The vegetation information from PLSRs is represented as discrete objects or continuous fields. Most PLSR studies, similar to the majority of vegetation mapping, employ discrete objects to express the presence or absence of individual species (Siccama, 1971; Grimm, 1984; Barrett et al., 1995; Dyer, 2001), or the dominant community type (Howell and Kucera, 1956; Loeb, 1987; Leitner et al., 1991). These expressions may be mapped as points (Anderson and Anderson, 1975; Brugam and Patterson, 1996) or lines (Batek et al., 1999), which in turn may be generalized to forest polygons using subjective boundaries based on apparently homogeneous patterns of abundant species (Siccama, 1971), or the point-in-polygon approach such as the association with mapped soil-type polygons (Veatch, 1925) or physiography (Liegel, 1982). These maps, however, are often manually created. The advent of Geographical Information Systems (GIS) has enabled the vegetation distribution to be automatically analyzed and visualized (Nelson et al., 1994; Barrett et al., 1995; Cogbill et al., 2002). In conjunction with GIS, programs have further been developed to derive geographic coordinates of bearing trees from the distances and directions recorded in PLSRs based on the topological characteristic of connectivity (Brown, 1998a; Batek et al., 1999). Public land survey digital base maps of section grids obtained from the U.S. Geological Survey Public Land Survey System or the Department of Natural Resources for each state are used as a template for the topological linking. The use of GIS allows efficient characterization and analysis of PLSRs which may overcome the obstacle of the trade-off between sample adequacy and data processing time (He et al., 2000).

Recent research in GIS and digital mapping also advanced methods for representing presettlement vegetation as continuous fields (Brown, 1998b). As opposed to the traditional discrete object representation, which is similar to the organism concept (Clements, 1936) and the community-unit concept (Shipley and Keddy, 1987) in ecology, the continuous field representation is founded on the continuum theory in ecology (Gleason, 1917) and the field view in GIScience (Burrough, 1996). Continuous fields are richer portrayals of vegetation distribution than discrete objects because they can illustrate gradual transitions, or ecotones, between vegetation types (Brown, 1998a). Presettlement vegetation distribution has been represented as continuous fields using interpolation methods and fuzzy sets (Brown, 1998a, 1998b; He et al., 2000).

C b. Investigation of vegetation and site relationships

Insights into the presettlement vegetation and site relationships are important because successful restoration is often based on matching species with their appropriate site requirements (niche) (Whitney and DeCant, 2001). Investigating the vegetation in relation to soils is one of the earliest utilizations of PLSRs (Veatch, 1925). In this mode of analysis, the spatial characteristic of bearing tree locations is analyzed to obtain vegetation-site relationships by locating trees on modern topographic or soil maps based on the assumption that site conditions of pedology, topography, and geology have not changed significantly since presettlement.

Earliest examination of vegetation-site relationships was conducted by descriptively summarizing and comparing species or forest type occurrences on various soil types or site conditions (Veatch, 1925; McIntosh, 1962). A more quantitative analysis was made by using multiple regression (Crankshaw et al., 1965). The introduction of binary discriminate analysis (also called contingency table analysis) by Strahler (1978) was a milestone in this mode of analysis. PLSR researchers, similar to that of other physical geographers, have taken full advantage of this method to investigate vegetation-site relationships (Whitney, 1982, 1990; Leitner et al., 1991; Abrams and McCay, 1996; Dyer, 2001). Associations between presettlement vegetation and site factors have also been developed using detrended canonical correspondence analysis (Cowell, 1995). Not until the late 1990s did PLSR researchers consider whether the size of the study area would affect the effectiveness of contingency table analysis. It has been suggested that at least 16 GLO townships must be included in a study so that most of the cells of the contingency table cells will be filled with a number other than zero (Almendinger, 1997).

The site relations of tree species have been investigated with respect to soil drainage and texture, topographic position, aspect, and slope, as well as surface geology and elevation (Lindsey, 1961; Siccama, 1971; Whitney and Steiger, 1985; Whitney, 1986; Barrett et al., 1995; Cowell, 1995; Abrams and McCay, 1996; Dyer, 2001; Whitney and DeCant, 2001). On the basis of analysis, the relationships between species and site conditions can be obtained. For example, in Baraga County, Michigan, jack pine was found to be restricted to the Baraga Outwash Plain (Barrett et al., 1995). Important environmental factors responsible for the presettlement vegetation distribution can also be determined. For example, soil drainage importantly controlled the Big Woods vegetation of Minnesota (Grimm, 1984), while topographic factors most strongly controlled species occurrence in southeastern Ohio (Dyer, 2001). It is, however, important to understand that use of forest types recreated using underlying geology and soils information in investigating vegetation-site relationships is problematic because the vegetation and site conditions are not independent anymore (Manies and Mladenoff, 2000).

C c. Identification of presettlement landscape patterns across a range of resolutions

Spatial resolution may greatly influence the detection and characterization of vegetation patterns, an issue related to the scale issue examined in depth in landscape ecology (Turner and Gardner, 1991; Qi and Wu, 1996). The PLSRs of the rectangular survey systems have multiple resolutions, such as the 9.6 × 9.6 km townships, the 1.6 × 1.6 km sections, and the 0.8 × 0.8 km quarter sections of the GLO surveys. Instead of attempting to determine the resolution of analysis for a given extent that is optimal for adequately quantifying presettlement vegetation patterns, PLSR researchers in the past simply used the data they obtained for the areas of their interest. A considerable amount of work done on the detection and characterization of spatial patterns in landscape ecological studies since the 1980s has then prompted researchers to investigate the effects of changing resolutions on pattern analysis with respect to the PLSRs.

In this mode of analysis, the spatial characteristic of different resolutions is analyzed to examine the influences on the resulting presettlement vegetation patterns. Studies of this sort can be divided into two types. The first type of study compares the reconstructions from different data sources at different sample resolutions. For example, witness trees recorded at lot corners were compared with the more finely-resolved line descriptions along the lot boundaries (Marks and Gardescu, 1992), and witness trees at the section-level were compared with the more finely-resolved witness trees at the lot-level (Mendelson, 2002). In both cases, the finely-resolved data indicate greater species richness and landscape heterogeneity than the coarsely-resolved data. The differences of the results from different resolutions may be caused by vegetation dynamics because the survey times of different resolutions were not the same. Coarser land division such as townships was conducted first, and finer subdivision of sections or lots was later.

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

The second type of study resamples the point tree data in the GLO records to assess how changes in the spatial resolution influence the apparent presettlement landscape heterogeneity. In general, landscape heterogeneity decreased as the spatial resolution coarsened. It was suggested that witness trees at quarter section corners in public land surveys were inadequate to capture the complexity of the presettlement landscape below several townships (Manies and Mladenoff, 2000). GLO data from Michigan, a resolution of one mile by one mile (one section), have been used to give a conservative characterization of landscape heterogeneity (Delcourt and Delcourt, 1996). Vegetation indices, including relative density, dominance, and importance values for each species, were calculated using the GLO data of Wisconsin at different resolutions, and the results suggested that thresholds of species heterogeneity were captured at a grid size of two sections by two sections (He et al., 2000).

B 2. Analysis of thematic characteristics

C a. Reconstruction of presettlement vegetation composition and structure

Concepts of presettlement vegetation composition and structure communicate and analyze aggregate characteristics of the data shown on the presettlement vegetation maps. Reconstruction of presettlement vegetation composition and structure are thus often accompanied with the creation of presettlement vegetation maps. This involves analyzing the thematic characteristics of species names, and distance and diameter measurements. Results of the composition and structure measures are often tabulated or graphed.

Description of presettlement vegetation composition is one of the major and earliest modes of geographically analyzing PLSRs. It can be derived from point tree data or line descriptions. Point tree data are most often used because they are more easily analyzed. Vegetation composition may be expressed as absolute numbers of witness/bearing trees for each species (Spurr, 1951) or as relative species percentages (Shanks, 1953; Howell and Kucera, 1956; Russell, 1981; Loeb, 1987; Cogbill, 2000). Line descriptions have been used in a few later studies (Seischab, 1990, 1992). They are believed to provide less biased information and a larger number of species than point tree data because surveyors simply recorded the species they encountered along the survey lines, and did not have to decide what trees to select as bearing trees, which might have resulted in biased representation and was noted as a concern in some studies (Bourdo, 1956; Grimm, 1984). Based on the assumption that species recorded in line descriptions were arranged in sequence of their abundance, as is known to be true in most GLO surveys (White, 1984), the line descriptions were quantified to derive forest composition in western New York (Seischab, 1992).

The compositional studies can be divided into three stages: understanding, comparing, and testing presettlement vegetation composition. Most studies have focused on understanding the vegetation composition in areas of their interests, especially those conducted in the early and mid-20th centuries (Lutz, 1930). Once the reconstructed presettlement vegetation composition is known, it can be compared with that of other adjacent regions or of different physiographic regions within the same study area (McIntosh, 1962; Seischab, 1992). Studies since the 1980s have further used the results to test and determine whether certain vegetation types (e.g., maple-basswood in Minnesota, oak savannas in Wisconsin) existed historically (Grimm, 1984; Schwartz, 1994; Dyer and Baird, 1997; Radeloff et al., 1998).

Presettlement vegetation structure can be studied by analyzing tree density and diameter class data from bearing tree records. This kind of study did not prevail until Cottam and Curtis (1956) introduced the idea of using distance between trees to obtain forest density. Tree density was then able to be calculated using distance measurements between bearing trees and designated survey corners. Estimated diameter measurements were available in most GLO surveys, but not in the metes and bounds and the private land surveys. Diameter classes, thus, can only be reconstructed in areas of GLO surveys. They have been compared among different vegetation types, and shown to approximate a reverse "J" shaped curve with the exception of the first size class in the forests of southern Illinois (Anderson and Anderson, 1975) and west-central Wisconsin (Schulte and Barnes, 1996). The lack of small trees (i.e., first size class) may be due to surveyor bias toward mid-size trees (Bourdo, 1956). Studies have further employed statistical tests to determine whether significant differences of vegetation structure occurred in different physiographic sections (Nelson, 1997).

Once the vegetation composition (i.e., frequency of occurrence) and structure (i.e., density) are obtained, importance values for individual species can then been calculated (Wuenscher and Valiunas, 1967; Delcourt and Delcourt, 1974; Anderson and Anderson, 1975; Nelson, 1997). It is suggested, however, that the density and diameters calculated from bearing trees might not be valid to assess vegetation structure due to the biased selection (Almendinger, 1997). Cautions must thus be employed when drawing conclusions based on the structural information.

C b. Comparison of presettlement and contemporary vegetation

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

It is critical to know how far the current landscape differs from its natural state (Baker, 1995) so as to appreciate how much change humans have caused and to obtain knowledge for ecological restoration. Studies often use presettlement vegetation as an ecological basis to compare with contemporary vegetation. The comparison involves analyzing the spatial characteristic of species locations and the thematic characteristic of vegetation composition to some degree. It is placed under the category of analyzing thematic characteristics in this paper because most studies focus on a given fixed area through which the vegetation composition change occurs. Studies of this type can be differentiated according to the available time periods for comparison, the source of data for contemporary vegetation, and the representation of the comparison.

There are typically two time periods used in the comparison, the presettlement and the present. In some cases data from intermediate time periods were available to provide insights into the transition: for instance, land use maps for the Illinois and Mississippi floodplain in 1903, 1935, and 1975 (Nelson et al., 1994), forest maps for the Allegany State Park region of New York ca. 1930 (Seischab, 1993), land economic surveys in Michigan ca. 1925 (Whitney, 1987) and Thoreau's account and survey in Concord, Massachusetts ca. 1850 (Whitney and Davis, 1986).

The source of data for contemporary vegetation includes field surveys and vegetation and land use maps. Firsthand field surveys were designed to approximate the presettlement land surveys, as in the quarter point method and random pairs method (Janke et al., 1978; Nelson et al., 1994; Van Deelen et al., 1996), or were obtained from transect or quadrat sampling (Cottam, 1949; Stearns, 1949; Cowell and Jackson, 2002). Secondhand field surveys were available from recent forest inventory databases such as the Ontario Forest Resource Inventory (FRI) (Jackson et al., 2000) and the USDA Forest Service Forest Inventory and Analysis (FIA) (Frelich, 1995; Dyer, 2001). These field surveys provide information on tree species, density and diameters, and are often combined with other secondhand field surveys such as inventories of old-growth forest (Frelich, 1995) to document changes of vegetation composition and structure. Caution, however, must be employed when comparing vegetation compositional estimates of presettlement forests from PLSRs with those from forest inventories because the PLSRs were spatially comprehensive with a small number of sampled trees, while the forest inventories were complete inventories for all the trees in a relatively small area (Shotola and Weaver, 1992). With the advent of remote sensing and GIS techniques, vegetation and land use maps of post-settlement and contemporary landscapes can be interpreted from satellite imagery or digitized using GIS (Iverson and Risser, 1987; Nelson et al., 1994; White and Mladenoff, 1994; Radeloff et al., 1999). The examination of landscape changes by comparing PLSRs to a satellite image classification, however, is not straightforward. The two datasets have very different data capture methods and spatial resolutions; thus pre-processing is required before comparison. For example, point tree data of the PLSRs have been interpolated to create grid layers (Radeloff et al., 1999), which is more interoperable with remote sensing imagery.

The representation of vegetation change over time can be expressed using a series of maps, summary tables, or the change detection approach. The first two representations, similar to the traditional ways to investigate vegetation change, have been used since the PLSR researchers started this mode of analysis. A series of maps of different snapshots have been used for graphical comparisons of tree species abundance (Radeloff et al., 1999) and landscape patch types (Whitney, 1987; White and Mladenoff, 1994). Summary tables have been used for numerical comparisons of areas and percentages of different vegetation communities or land use types (Cottam, 1949; Stearns, 1949; Siccama, 1971; Iverson and Risser, 1987; Frelich, 1995). In addition to the graphical and numerical comparisons, change detection, a common approach for investigating land-cover change in remote sensing (Lee and Lunetta, 1995), provides the PLSR researchers a new method to conduct the comparison. Transition matrices are often used in the change detection approach to show changes in tree species abundance between presettlement and the present (White and Mladenoff, 1994; Radeloff et al., 1999). Changes in land use or vegetation communities from one category to another can be illustrated in maps (Iverson, 1988; White and Mladenoff, 1994). The change detection approach thus offers more quantitative insights into the changes that have occurred than the other two representations. When using this approach, the locations of water bodies have often been assumed to be stable, and the distribution of current water bodies have been superimposed on the presettlement vegetation maps (Nelson et al., 1994; Radeloff et al., 1999).

After investigating the landscape changes from the presettlement to the present, researchers can then decipher the changes and predict future trends. Anthropogenic influences such as fire suppression, logging and clearing have been found to be important factors affecting landscape change (Whitney, 1987; Palik and Pregitzer, 1992; White and Mladenoff, 1994; Frelich, 1995; Radeloff et al., 1998; Cowell and Jackson, 2002). Natural disturbances such as windstorms, deer or moose browsing (Janke et al., 1978; Van Deelen et al., 1996) and pathogens (e.g., chestnut blight) (Seischab, 1993) have been considered responsible for the changes in some areas.

B 3. Analysis of temporal characteristics

C a. Study of disturbance regimes

It is now recognized that ecological preservation requires the maintenance of natural disturbance (Radeloff et al., 2000). In many cases the natural disturbance regime has been fundamentally changed by settlement and is therefore unknown. The temporal characteristic of age-related information recorded in PLSR line descriptions can be analyzed to understand disturbance regimes, as do dendrochronology and pollen studies. This mode of analysis did not appear until the late 1970s when Lorimer (1977) used PLSRs to calculate the natural disturbance cycle of northeastern Maine. It has mainly been conducted in the public GLO and private land survey areas because of the line descriptions they provided. The idea is to use line transect sampling theory to determine the total areas affected by fire or windthrow present at the time of the surveys and to estimate the rotation period (the time period in which fire or windthrow will disturb a spatial extent equivalent to the study area) by assuming that disturbances would only be observed for a certain number of years after their occurrences. Detailed calculating procedures were elucidated in Canham and Loucks (1984). Researchers have used the methodology to compare a specific disturbance regime in different regions based on vegetation type or topography (Whitney, 1986; Seischab and Orwig, 1991; Zhang et al., 1999), or to discuss different disturbance regimes such as windstorms and fires of a specific region (Lorimer, 1977).

The estimated disturbance return times have varied between different areas. Windthrow return times of 980 and 3190 years were determined for the Phelps and Gorham Purchase and the

Holland Company Lands in New York respectively although the two areas are adjacent to each other (Seischab and Orwig, 1991). In northern Maine, the average return time of fire and large-scale windthrow were estimated to be 800 and 1150 years respectively (Lorimer, 1977). It has also been estimated that fire reoccurrence can range from as short as 80 years for presettlement jack pine forests to as long as 1200 years for the hemlock-white pine-northern hardwoods forests in Lower Michigan (Whitney, 1986). In addition to measures of disturbance return times, studies have discussed the role of disturbance regimes, especially of fire, in shaping the presettlement vegetation pattern (Grimm, 1984; Whitney and Steiger, 1985; Cowell, 1995).

Just as the use of dendrochronology to characterize surface fire regimes is problematic because many trees are not scarred during a fire (Batek et al., 1999), the use of PLSRs in studying presettlement disturbance regimes also has its pitfalls. The line descriptions document only the more catastrophic events, and provide only a snapshot of the disturbance at the time of the survey and may overlook unusual climatic events conducive to windthrows or the spread of fires (Whitney, 1986; Zhang et al., 1999). Despite these limitations, the PLSRs still provide some of the best information on disturbance regimes, and have contributed significantly to the understanding of North American's presettlement ecosystems (Whitney and DeCant, 2001).

A V. Data Quality of the Presettlement Land Survey Records

The above section reviews how various geographic characteristics of PLSRs have been analyzed in previous studies. For the most appropriate future use of PLSRs, an understanding of the quality and limitations of the data is needed. The meaning of 'quality' depends on the context in

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

which it is applied (Veregin, 1999). The Data Quality Specification of the Spatial Data Transfer Standard (SDTS) defines data quality as information a producer should present to a potential user to allow the user to make a determination of fitness for some particular use (Chrisman, 1997). Reporting error is therefore not a sign of weakness because the error estimate provides crucial information which must be preserved for correct interpretation (Chrisman, 1991: 166). The five data quality components used in GIScience- lineage, positional accuracy, attribute accuracy, logical consistency, and completeness- are employed here to examine the qualities of the geographic characteristics of PLSRs. We can then be aware of the biases and limitations of the PLSRs for presettlement vegetation studies, comprehend whether these issues will potentially affect the accuracy of the results of such studies, and understand how to prevent or minimize potential problems by knowing how previous studies have tackled these issues.

B 1. Lineage

Lineage is a description of the source materials, methodologies (including the acquisition and compilation of the source data), and transformations used to produce a database (Johnston, 1998). For PLSRs, surveyor’s instructions provide information about lineage, especially on the acquisition of the source data. Researchers who refer to PLSRs must become acquainted with the survey procedures in their study areas (Bourdo, 1956). They can then understand what, how, and when the data were collected, and therefore envision the potential data quality issues of PLSRs. For example, knowing who collected the data will help spot poor surveys. Similarly, knowing how long the survey of an area took may help explain whether vegetation differences between adjacent townships resulted from differences in site conditions, anthropogenic factors of burning and harvesting (Schulte and Mladenoff, 2001), or even natural processes of climate change, which can result in ecosystem shift and can occur within a time scale of centuries (Davis, 1986). The metes and bounds surveys generally did not have instructions for surveyors. The GLO surveys did have instructions, but unfortunately they differed from place to place and time to time. In general, though, the instructions of 1855 guided practically all of the GLO surveys from then until the end of the contract system in 1910 (Dodds et al., 1943, from Bourdo, 1956).

In addition to the errors from the data acquisition process, the compilation process of original survey field notes to archived microfilms might contain errors in spatial characteristics such as bearing tree locations, thematic characteristics such as species names, and temporal characteristics such as time of observation. A comparison of the original field notes and the further removed GLO microfilms shows that there is approximately a 1-5% transcription error rate for corner records (Almendinger, 1997). It is suggested that errors occurred during the hand-copying procedure executed by the clerks in the Surveyor General's office or the transcription of the data from the GLO microfilms to data-entry forms by researchers.

Although the importance of lineage of PLSRs was recognized as early as in Bourdo's study (1956), the issues of data acquisition and composition have only been considered in a few studies in the past half century. Surveyors' instructions and reputations were inspected to understand the data acquisition process (Delcourt and Delcourt, 1974; Whitney, 1986). Spatial and thematic characteristics of GIS-generated presettlement vegetation maps were verified with original survey notes (Barrett et al., 1995) to reduce the errors resulted from data compilation. I here recommend that researchers look into surveyors' instructions if available before their studies, or explore the historical background of surveys when the surveyors' instructions are unable to be located. Quality control of the researchers' data compilation processes is also essential. Since the reliability of PLSRs has been concerned in some studies (Grimm, 1984), researchers should prevent any human-induced errors that might further affect the quality of the data. A guide to understanding and interpreting the GLO records has been prepared for the Illinois survey (Hutchison, 1988), but the suggested techniques and interpretations may be applied to other parts of the USA so researchers can better understand the notes and use them more efficiently.

B 2. Positional accuracy

Positional accuracy refers to the accuracy of the coordinates within the database (Johnston, 1998). It mainly influences the qualities of the spatial characteristics. Three positional accuracy issues of PLSRs that need to be considered are mislocation of corners, mislocation of landscape features, and incorrect positions for bearing trees.

C a. Mislocation of corners

Mislocation of corners influences the spatial sub-characteristic of location (Table 1, Figure 4). If a survey corner is given an incorrect location, it will then cause erroneous topological relationships and distorted geometry in surveyors' township plats. This type of problem sometimes occurred due to inadvertent or unavoidable errors. For example, iron-ore deposits in the Ozark Mountains of Missouri caused major deviations in magnetic compass readings that led to mislocation of survey corners and broken, disconnected, and straying survey lines (Corner et al., 2000: 47). Errors of this sort, however, do not necessarily affect the usefulness of PLSRs (Bourdo, 1956). For example, the mislocation of corners may not be a significant issue if the resulting point bearing tree data are used to create vegetation maps over a large area. It may be a critical issue, however, if the resulting township plats are used to investigate the vegetation change from presettlement to the present, especially in a small area of a township.

Figure 4 here#

C b. Mislocation of landscape features

Landscape features such as streams and wetlands may have been mislocated on GLO township plats. This issue influences the spatial sub-characteristic of location as does the issue of mislocation of corners. The GLO surveyors recorded what they encountered along the survey lines and then interpolated the boundaries of landscape features between the lines (Whitney and DeCant, 2001). Since surveyors did not visit every place in space, landscape features such as courses of streams and boundaries of wetlands inside survey lines might be incorrect (Figure 5), and areas of small landscape features such as relict pine stands might be underestimated. This issue is uncommonly addressed in PLSR studies, but has been resolved through calibration and rectification. Empirical calibration allows the underestimated presettlement wetlands to be corrected (Iversen and Risser, 1987). Rectification of basic hydrological features on plats by relating with USGS topographic maps and other geo-referenced aerial photos allows other types of land cover information on plats to be transferred, thus improving the accuracy of the GLO township plats (Bandi and Huggins, 1994).

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

Figure 5 here#

C c. Incorrect positions for bearing trees

The words 'adjacent' and 'nearly' in the surveyor's instructions and PLSRs do not necessarily imply that bearing trees were the closest individuals to a survey corner (Nelson, 1997). For example, the surveyor's instructions in 1831 and 1850 required surveyors to select the closest trees in each section, while the instructions of other years required that bearing trees only be in different sections without considering adjacency (Bourdo, 1956; White, 1984). In addition, even though the surveyor's instructions required surveyors to record the closest trees to each survey corner, they might have been concerned with tree size, longevity, and economic value (Grimm, 1984). These surveyor biases may have resulted in bearing trees not being the closest trees to the survey corners, therefore influencing the spatial sub-characteristic of topology. In other words, the locations of the recorded bearing trees may not represent the real bearing tree positions. As a consequence, this positional accuracy issue influences the attribute accuracy of distance measurement, which will be discussed in the next section.

This positional accuracy issue resulted from biases in bearing tree selection is often addressed in PLSR studies. Researchers have estimated biases in their data using statistical techniques such as analysis of mean post-to-tree distance (Bourdo, 1956), chi-square test (Bourdo, 1956; McIntosh, 1962), analysis of variance (ANOVA) (Delcourt and Delcourt, 1974; Dyer, 2001), and Spearman rank correlation test (Manies et al., 2001). Although some have argued that so many statistical assumptions are violated that the tests are unreliable (Grimm, 1984), other researchers have suggested that previous statistical analyses are appropriate (Maines et al., 2001), and biases in bearing tree selection are not serious concerns especially over large areas (Almendinger, 1997).

B 3. Attribute accuracy

Attribute accuracy refers to how accurately the thematic characteristics have been identified, classified, and quantified (Johnston, 1998). Identification and classification accuracy occurs for nominal and ordinal levels of measurement, while quantification accuracy occurs for ratio and cyclical levels of measurement.

C a. Erroneous identification of tree species

Erroneous identification of tree species relates to the quality of the nominal measurement of PLSRs. It may result from intentionally fraudulent surveys or unintentional surveyors' misidentifications. Fraudulent surveys, which were manufactured by surveyors without going into the field, have been noted in some GLO survey areas, but were corrected by resurveys in most cases (Bourdo, 1956). Surveyors' misidentifications are of concern because the surveys were not conducted for ecological purposes; therefore, it has been questioned whether the surveyors were able to accurately identify tree species (Forman and Russell, 1983), and thus whether the PLSRs can represent the true presettlement landscape. Studies have argued, however, that based on the wide variety of tree species noted in many surveys, the surveyors generally had a good knowledge of vegetation identification (Spurr, 1951; Whitney, 1996).

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

C b. Semantic heterogeneity in tree species names and soil quality classification

Semantic heterogeneity in tree species names influences the quality of the nominal measurement, and thus hampers the reconstruction of presettlement vegetation to the species level. This issue results from the use of common tree names, which vary from region to region and from surveyor to surveyor (Whitney and DeCant, 2001). For example, some surveyors used 'soft maple' to include both silver maple (*Acer saccharinum* L.) and red maple (*A. rubrum* L.) (Seischab, 1992). Other collective names such as hickory, oak, or birch may indicate one or more species (Lutz, 1930). For example, the GLO surveyors in southwestern Illinois did not differentiate hickory and oak species which are important in localized areas (Brugam and Patterson, 1996). The typical approach with such species has been to conservatively group them to the genus level (Finley, 1976). This is the most often addressed data quality issue in PLSR studies. Attempts have been made to qualitatively interpret surveyors' common tree names in terms of the current scientific nomenclature using early taxonomic guides or based on species' current distributions (Lutz, 1930; Shanks, 1953; Schafale and Harcombe, 1983; Grimm, 1984; Fralish et al., 1991; Seischab, 1992; Barrett et al., 1995; Almendinger, 1997; Dyer, 2001). A more quantitative analysis of logistic regression models has recently been developed to narrow the uncertainty in species classification and has demonstrated potential to increase the usefulness of PLSRs (Mladenoff et al., 2002). Thus, methods to examine this issue include interpreting tree type equivalents, developing logistic regression models, or using additional information from PLSRs which will be discussed in the next section of logical consistency. The generalization of species data to the genus level should be avoided because ecological information may be lost, especially where

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

congeneric species differ ecologically and where a relatively high proportion of individual trees are undifferentiated (Mladenoff et al., 2002). Nonetheless, it is important to note that this approach may not affect the comparison of presettlement and modern vegetation because some of the current forest inventories (e.g., FIA) often do not differentiate species that are difficult to distinguish (e.g., hickories).

Another semantic heterogeneity issue is classification of the ordinal measurement of soil quality. The ‘second rate’ soil in the Midwest may indicate something quite different from the ‘second rate’ soil in the arid West (Galatowitsch, 1990). On the other hand, different ordinal measurements may refer to the same condition, for example, ‘well-drained’ soil and land of ‘the first quality.’ Prior studies have not employed the PLSRs in reconstructing soils or site conditions as it is often assumed that they have remained stable since the presettlement era. This issue, however, potentially impedes the use of PLSRs in comparing or integrating soil or land quality with different regions.

C c. Imprecise quantifications of distance, size, and bearing

The qualities of distance and size measurements influence the accuracy of vegetation structure reconstruction because distances from the closest trees to the corners, which are used to calculate forest density, and tree sizes, which are used to obtain size-class structure, will be in error. Distance measurements vary in quality because, although the survey unit of chains and links was used in most PLSRs (There are 100 links in one chain), some metes and bounds surveyors tended to use five-link increments to record distance (McIntosh, 1962). The private and public land surveys, however, had more precise distance measurements. The GLO surveyor's instructions even requested that cautions be made when measuring distances at non-level locations such as hilly regions (White, 1984). Size measurements vary in quality because trees below a convenient marking size were ignored. For example, some GLO surveyors favored mid-size trees ranging from 25.4 to 35.56 cm (10 to 14 in) DBH (Bourdo, 1956; Schafale and Harcombe, 1983; Grimm, 1984). The GLO surveyors in Minnesota tended to use even-inch measurements to record DBH (Almendinger, 1997). These issues, as the positional accuracy issue of incorrect positions for bearing trees, may result from biases in bearing tree selection and are often addressed in PLSR studies. Grimm (1984) and Almendinger (1997) argued that biases in bearing tree selection do not permit quantitative reconstruction of absolute tree density or tree size. However, the data do still provide adequate information on relative structure of woods (Whitney and DeCant, 2001) such as prairie vs. oak savanna, especially when complemented by line descriptions (Schulte and Barnes, 1996).

Cyclical measurements can vary in quality, as in some metes and bounds surveys where only the cardinal bearings of north and south were used (McIntosh, 1962). The low quality in quantifying these attribute measures influences the positional accuracy by making the precise relocation of bearing trees impossible. The data compilation process might also create or exacerbate this problem. For example, earlier compilation of the Minnesota bearing tree database used imprecise bearings (e.g., NE rather than N37°E) although the precise information was available from surveyors' field notes. This issue will have minimal effects on creating presettlement vegetation maps if PLSRs are used at a landscape extent.

B 4. Logical consistency

Logical consistency refers to the internal consistency of the rules of representation. It verifies relationships that should be present (Chrisman, 1997) such as valid values for thematic characteristics and consistent spatial locations for the same landscape features derived from different data sources. If logical consistency does not occur, potential errors are implied. It can thus be used to examine the qualities of the spatial and thematic characteristics of PLSRs.

For the spatial characteristics, if landscape features such as rivers and streams are recorded in line descriptions then they should also be found at similar locations in township plats. In addition, logical consistency of topological characteristics such as adjacency and connectivity requires that land parcels next to each other share a non-overlapping boundary. The recorded line descriptions of the north boundary of township Number One should thus be the same as the south boundary of township Number Two, which is on the north side of township Number One in the same range. The rectangular GLO and private land surveys usually were topologically consistent. The irregular metes and bounds surveys, however, had more topological inconsistencies, resulting in many cases of overlap (White, 1984).

For the thematic characteristics, logical consistency can be used to solve the attribute accuracy issue of semantic heterogeneity in tree species names, and to investigate whether biases existed in bearing tree selection and whether the ratio measurement of distance had errors. For example, we can determine whether a maple noted in the bearing tree data was a sugar maple (*A. saccharum* Marsh.), red maple (*A. rubrum L.*), or silver maple (*A. saccharinum* L.) by using auxiliary information such as its association with other species mentioned in line descriptions, township plats, or current vegetation distribution maps. If a maple in a wetland situation was found in association with elm and black ash it would be interpreted as silver maple (*A. saccharinum* L.), but if it was found in association with cedar or larch it would be interpreted as red maple (*A. rubrum* L.) (Seischab, 1992). Biases in bearing tree selection can be investigated by comparing the relative frequencies of tree types as bearing trees vs. line trees based on the assumption that results from the two sources of point tree data should be consistent. When inconsistencies occur, the PLSR researchers assume that the relative frequencies derived from line trees provides a more accurate estimate than those derived from bearing trees because line trees are free from biases associated with the task of blazing and scribing certain kinds of trees as bearing trees (Almendinger, 1997). As for the ratio measurement of distance, the survey unit of one link, which is equal to 20.1168 cm, was usually adequate for recording the distances between the bearing trees and the designated survey corners in forested areas. If the survey unit of one chain, which is equal to 100 links, was found in this case, it might imply an erroneous distance measurement or a data-entry error.

B 5. Completeness

Completeness refers to the degree to which the database exhausts the universal set of various features (Johnston, 1998), and is applicable to examining the qualities of the spatial and thematic characteristics. The completeness issues of PLSRs include incomplete area coverage, biased representation, unavailable geometric features, and incomplete species lists.

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

C a. Incomplete area coverage

The metes and bounds survey did not systematically survey the land due to the irregular nature of the survey system which led to many land gaps. Also, the regular survey systems of the GLO and the private land surveys did not cover some areas such as Indian reserves (Figure 3). This issue affects the qualities of the spatial sub-characteristics of location and resolution because it reduces the sample size of bearing trees for analysis and impedes the reconstruction of presettlement vegetation in areas where the PLSRs were not covered. The use of the data at a landscape extent may reduce the effect of this issue.

C b. Biased representation

The number of sampled trees and the type of land represented in surveys may be biased against difficult terrain. Areas of complex topography were not surveyed completely by surveyors (McIntosh, 1962; Liegel, 1982; Manies and Mladenoff, 2000; Black and Abrams, 2001). For example, in the metes and bounds surveys, higher altitudes had sparser observations. The surveyors' records were most representative of the forest cover of the valley areas and the slopes below 914.4 m (3000 ft), while there was inadequate representation of mountaintops and higher ridges which rise to 1219.2 m (4000 ft) (McIntosh, 1962). The study using the Wisconsin GLO records showed that inadequate representation occurred where topography was complex (Liegel, 1982). The issue of biased representation in rough topography, as does the issue of incomplete area coverage, influences the qualities of the spatial sub-characteristics of location and resolution.

C c. Unavailable geometric features

The geometric features of point, line, and polygon that are the abstractions of bearing trees, line descriptions, and township plats, are not all available in the three types of PLSRs as discussed in the section on geographic characteristics, and thus affect the quality of the spatial characteristic of geometry. The metes and bounds surveys often contained the point geometric feature but not the other two geometric features. The private HLC survey included both the point and line geometric features, and a few sketch maps along the survey lines provided the polygon geometric feature. The GLO surveys contained all three geometric features. Different geometric features offer different types of information on the presettlement landscape. For example, the point features of bearing trees are used most often in recreating presettlement forest patterns, compositions, and structure. The occurrence of natural disturbance recorded in line descriptions is used to reconstruct presettlement disturbance regimes, so the quality of the line features further affects the quality of the temporal characteristic. The polygon features of wetlands and prairies on township plats are used in landscape change studies of small extent. The lack of certain geometric features, therefore, constrains the application of PLSRs in certain modes of analysis. Researchers need to consider their specific study objectives and the availability of geometric features in their PLSRs so as to select the most appropriate feature for analyses.

C d. Incomplete species list

Because the surveyors did not visit every location in the presettlement landscape, one can argue that no PLSR could provide a complete species list, a data quality issue related to the thematic characteristic of nominal measurement. Indeed, plant ecologists caution against the use of regular survey systems because they can miss small communities and, more importantly, community types regularly distributed but off the survey line (Barbour et al., 1999). Nonetheless, the regular GLO and the private land surveys might provide a more complete species list than the irregular metes and bounds surveys because they included line descriptions, in which surveyors could list as many tree types as they wanted.

These incompleteness issues, as well as the above logical inconsistency issues, received less attention from prior researchers. A reason for this may be that studies examining the limitations of PLSRs only focus on the GLO surveys, which did not have as many quality issues regarding logical consistency and completeness as the earlier metes and bounds and private land surveys did. Among the other data quality issues of lineage, positional accuracy, and attribute accuracy, researchers usually address one or two recognized issues in their PLSRs, with the positional accuracy issue of incorrect positions for bearing trees and the attribute accuracy issue of semantic heterogeneity in tree species most often addressed. The GIScience approach of this study shows that the data quality issues that need to be addressed are not just those that have received the most attention. Also, some quality issues are more important with certain PLSR types. For example, the metes and bounds surveys usually have issues of logical consistency and completeness, while the GLO surveys don't. Depending on the purposes and the spatial extent of the research, the potential data quality issues may or may not have an influence on the best approach. In the case of reconstructing presettlement vegetation of an entire state, the limitations of incorrect positions for bearing trees and underestimation of species richness may not affect the accuracy of the results. The use of the five data quality components from GIScience therefore provides a more comprehensive and systematic analysis of the data quality issues of all three types of PLSRs than previous studies.

A VI. Conclusions and Future Directions

Presettlement land survey records (PLSRs) provide unique and valuable information regarding landscape patterns prior to European disturbance. These data have been suggested to be useful in understanding presettlement vegetation conditions and human impacts on the landscape. The usefulness of the data, however, varies with the characteristics and qualities of the data. In this study, a theoretical framework from GIScience is employed to examine the geographic characteristics of PLSRs, to review PLSR studies based on the geographic characteristic they analyzed, and to investigate quality issues of the data.

The PLSRs can be classified into three types: metes and bounds surveys, public GLO surveys, and private land surveys. The main geographic characteristics used to describe PLSRs are space, theme, and time, each of which is further divided into sub-characteristics. The three types of PLSRs are found to have distinctly different geographic characteristics, especially the spatial sub-characteristics of shape and size/resolution. In addition, the significances of the geographic characteristics for presettlement vegetation studies are discussed.

A review of previous PLSR studies identified six major modes of analysis, which can be classified according to the three characteristics of space, theme, and time. Modes of analyzing spatial characteristics have included creation of presettlement vegetation maps, investigation of vegetation and site relationships, and identification of presettlement landscape patterns across a range of resolutions. Modes of analyzing thematic characteristics have consisted of reconstruction of presettlement vegetation composition and structure and comparison of presettlement and contemporary vegetation. The only mode of analyzing temporal characteristics identified is the study of disturbance regimes. The trend in each mode of analysis is reviewed with newly developed theories and methodologies in geography and ecology, such as the advent of GIS and remote sensing techniques for examining vegetation dynamics and the introduction of contingency table analysis for determining vegetation-site relationships. Furthermore, cautions and recommendations are provided for appropriate utilization of the geographic characteristics analyzed in each mode of analysis.

Data qualities of the geographic characteristics of PLSRs are examined using the data quality components of GIScience, including lineage, positional accuracy, attribute accuracy, logical consistency, and completeness. Additionally, methodologies which previous studies have employed to improve the quality and usefulness of the data are summarized. As Chrisman (1991: 166) mentioned in a discussion of error and quality in spatial data, “Reporting error is not a sign of weakness… because the error estimate provides crucial information which must be preserved for correct interpretation.” In other words, the data quality issues of PLSRs do not render data useless; understanding these issues can help researchers realize the usefulness and limitations of PLSRs, and then make better utilization of the data. Depending on the purpose and the spatial extent of a researcher’s interests, the potential data quality issues identified in this paper may or may not have an influence on a particular study. It is suggested that many quality issues, such as mislocation of corners, incorrect positions for bearing trees, and incomplete area coverage, may be alleviated when the PLSRs are used for analysis over a large spatial extent.

This study demonstrates not only that GIS techniques are useful in presettlement vegetation studies, but also that a GIScience framework can provide a systematic organization for reviewing the geographic characteristics and data quality issues of PLSRs, and studies that have used the data. Knowledge of these geographic characteristics is important so researchers can be aware of the characteristics of the data and thus the modes of analysis open to them. For example, private HLC survey records lack the thematic characteristic of diameter measurements; therefore, it will be difficult to use the data to reconstruct presettlement vegetation structure. Insights into the data quality issues are essential so researchers can realize the potential limitations of the geographic characteristic they utilize. For example, although public GLO survey records have diameter measurements, they can only be used for relative rather than absolute vegetation structural reconstruction due to the issue of imprecise quantification of size measurement that might be attributed to biases in bearing tree selection. The systematic examination of this paper thus provides a more comprehensive understanding of the usefulness and limitations of PLSRs than previous studies, and can assist future studies to employ the PLSRs most appropriately. Researchers can then avoid invalid analyses and erroneous conclusions. Furthermore, it can be served as a framework to review PLSRs in Australia and other sources of historical data in geography and ecology.

In the future, the following possible research directions for different modes of analysis deserve special attention. For the creation of presettlement vegetation maps, instead of creating boundaries between different vegetation types using existing boundaries of soils or topography, spatial models that incorporate environmental variables can be developed in concert with GIS and spatial interpolation methods to generate continuous maps of species and forest types using point tree data. Scant attention has been paid to the use of line descriptions, with the notable exception of Seischab (1990, 1992). Since surveyors could list as many tree types as they wanted, surveyors' biases would least likely happen in line descriptions as opposed to point bearing trees. Therefore, another possible research topic for creating presettlement vegetation maps would be to incorporate information from line descriptions, specifically on quantifying the species recorded in the line description if they were arranged in order of their abundance. The results would also contribute to a more accurate reconstruction of presettlement vegetation composition.

Studies that engage in investigating presettlement vegetation-site relationships can evaluate whether such relationships have changed following forest clearing and subsequent regrowth (Dyer, 2002). Because most studies have used the current site conditions in their investigations, it is also important to know whether the sites with different soils and geological data, digital terrain models, and scales of analysis influence the results. The PLSRs provide a valuable source of information on site-specific determinants of presettlement ecosystems that are important for successful restoration. Only a few studies, however, have used this information for ecosystem restoration activities such as relocating savanna and barren sites (Stritch, 1990). Restoration ecologists thus should take advantage of the outcomes from the vegetation-site investigations. On the other hand, they should also keep in mind that the resulting relationships are relatively closer not absolutely equivalent to the potential natural conditions.

The use of PLSRs in analyses of larger spatial extents, such as those of several counties or a state, may alleviate some of the limitations of the data. To map large-area presettlement landscapes, it is desirable to know whether the coarsely-resolved township data can predict landscape patterns similar to those that would be predicted by the finely-resolved section data. The reasons are twofold. First, the amount of time required for data entry is much less with the township data than with the section data. Second, the township-level data provides a more temporally precise picture of the presettlement landscape than the information from the section-level data because the township boundary surveys were usually conducted first and accomplished within a few years, as opposed to the time span of several decades of the section subdivision surveys. One possible approach would be to develop a scheme to generate bearing tree data at resolutions that can represent the township-level and section-level data respectively, and to identify how the reconstructed presettlement landscape patterns change with respect to those resolutions. The results would be useful for efficiently creating more temporally precise maps of presettlement vegetation of a state or several states.

Vegetation patterns and compositions reconstructed from PLSRs are often compared with the contemporary vegetation to examine vegetation dynamics. The PLSRs and current forest inventories, however, have different sampling schemes. The PLSRs are dense network of sample locations, but fewer trees sampled per location; current forest inventories are sparse network of sample locations, but more trees sampled per location. Statistical methods need to be developed that can incorporate the two different sampling schemes to achieve a more accurate comparison of the landscapes they came from. For studies of disturbance regimes of fires, the PLSRs should be used in conjunction with pollen analyses, charcoal records, or tree-ring data (Schulte and

Mladenoff, 2001). Recent research has demonstrated the integration of dendrochronology with the disturbance information derived from PLSRs (Batek et al., 1999). The limitation that PLSRs provide only a snapshot of disturbance history can be supplemented by the advantage of dendrochronology of providing specific, quantifiable information on disturbance regimes and how they varied across the landscape.

There has been a great amount of PLSR papers published since the early 20th century. Most studies deal with areas of the eastern USA, especially in the Northeast and Great Lakes areas. Compared with the vast extent of the lands surveyed, the use of PLSRs in presettlement vegetation studies is still a field of great potential.

References:

Abrams, M.D. and McCay, D.M. 1996: Vegetation-site relationships of witness trees (1780-1856) in the presettlement forests of eastern West Virginia. Canadian Journal of Forest Research 26, 217-224.

Almendinger, J. 1997: Minnesota's bearing tree database. Biological Report 56, St. Paul, MN: Minnesota Department of Natural Resources.

Anderson, R.C. and Anderson, M.R. 1975: The presettlement vegetation of Williamson County, Illinois. Castanea 40, 345-363.

Avery, T.E. 1967: Forest measurements. New York, NY: McGraw-Hill.

Baker, W.L. 1995: Long-term response of disturbance landscapes to human intervention and global change. Landscape Ecology 10, 143-159.

Bandi, D.E. and Huggins, D.G. 1994: Converting public land survey information into digital maps of improved accurace and usefulness. In Congalton, R.G., editor, Proceedings of the 1st International Symposium on the Spatial Accuracy of Natural Resource Data Bases, Bethesda, MD: American Society for Photogrammetry and Remote Sensing, 35-46.

Barbour, M.G., Burk, J.H., Pitts, W.D., Gilliam, F.S. and Schwartz, M.W. 1999: Terrestrial plant ecology, third edition. Menlo Park, CA: Addison Wesley Longman Inc.

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

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

Black, B.A. and Abrams, M.D. 2001: Influences of Native Americans and surveyor biases on metes and bounds witness-tree distribution. Ecology 82, 2574-2586.

Bonham-Carter, G.F. 1994: Geographic information systems for geoscientists: modelling with GIS. Pergamon, NY: Oxford Press.

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.

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.

Brugam R.B. and Patterson, M.J. 1996: Application of a geographic information system to mapping presettlement vegetation in southwestern Illinois. Transactions of the Illinois State Academy of Science 89, 125-141.

Bürgi, M. and Russell, E.W.B. 2000: Evaluating accuracy of two types of early land survey records in the northeastern United States. Journal of the Torrey Botanical Society 127, 94-98.

Burrough, P.A. 1996: Natural objects with indeterminate boundaries. In Burrough, P.A. and Frank, A.U., editors, Geographic Objects with Indeterminate Boundaries, London, UK: Taylor & Francis Ltd, 3-28.

Canham, C.D. and Loucks, O.L. 1984: Catastrophic windthrow in the presettlement forests of Wisconsin. Ecology 65, 803-809.

Chrisman, N.R. 1991: The error component in spatial data. In Maguire, D.J., Goodchild, M.F. and Rhind, D.W., editors, Geographical information systems: principles and applications, New York, NY: John Wiley and Sons Inc., 165-174.

Chrisman, N.R. 1997: Exploring geographic information systems. New York, NY: John Wiley and Sons Inc.

Clements, F.E. 1936: Nature and structure of the climax. Journal of Ecology 24, 252-284.

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. and 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.

Corner, J., Maclean, A.S. and Van Valkenburgh, M. 2000: Taking Measures Across the American Landscape. Cambridge, MA: Yale University Press.

Cottam, G. 1949: The phytosociology of an oak woods in southwestern Wisconsin. Ecology 30, 271-287.

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

Cowell, C.M. and Jackson, M.T. 2002: Vegetation change in a forest remnant of the eastern presettlement prairie margin, USA. Natural Areas Journal 22(1), 53-60.

Cowell, C.M. 1995: Presettlement piedmont forests: patterns of composition and disturbance in central Georgia. Annals of the Association of American Geographers 85, 65- 83.

Crankshaw, W.B., Qadir, S.A. and Lindsey, A.A. 1965: Edaphic controls of tree species in presettlement Indiana. Ecology 46, 688-698.

Davis, M.B. 1986: Climatic instability, time lags, and community disequilibrium. In Diamond, J. and Case, T.J., editors, Community ecology, New York, NY: Harper & Row Publishers, 269-287.

Delcourt, H.R. and Delcourt, P.A. 1974: Primeval magnolia-holly-beech climax in Louisiana. Ecology 55, 638-644.

Delcourt, H.R. and Delcourt, P.A. 1977: Presettlement magnolia-beech climax of the Gulf Coastal Plain: quantitative evidence from the Apalachicola River bluffs, north-central Florida. Ecology 58, 1085-1093.

Delcourt, H.R. and Delcourt, P.A. 1996: Presettlement landscape heterogeneity: evaluating grain of resolution using General Land Office Survey data. Landscape Ecology 11, 363-381.

Dyer, J.M. and Baird, P.R. 1997: Remnant forest stands at a prairie ecotone site: presettlement history and comparison with other maple-basswood stands. Physical Geography 18, 146-159.

Dyer, J.M. 2001: Using witness trees to assess forest change in southeastern Ohio. Canadian Journal of Forest Research 31, 1708-1718.

Dyer, J.M. 2002: Using digital environmental data with field sampling to evaluate presettlement species-site relationships. 98th Annual meeting of the Association of American Geographer, Los Angeles, CA.

Fensham, R.J. and Fairfax, R.J. 1997: The use of the land survey record to reconstruct pre-European vegetation patterns in the Darling Downs, Queensland, Australia. Journal of Biogeography 24, 827-836.

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

Forman, R.T.T. and Godron, M. 1986: Landscape ecology. New York, NY: John Wiley and Sons Inc.

Forman, R.T.T. and Russell, E.W.B. 1983: Evaluation of historical data in ecology. Bulletin of the Ecological Society of America 64, 5-7.

Foster. D.R. 1992: Land-use history (1730-1990) and vegetation dynamics in central New England, USA. Journal of Ecology 80, 753-772.

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

Frelich, L.E. 1995: Old forest in the Lake States today and before European settlement. Natural Areas Journal 15(2):157-167.

Galatowitsch, S.M. 1990: Using the original land survey notes to reconstruct presettlement landscapes in the American West. Great Basin Naturalist 50, 181-191.

Gleason, H.A. 1917: The structure and development of the plant association. Bulletin of the Torrey Botanical Club 44, 463-481.

Gordon, R.B. 1940: The primeval forest types of southwestern New York. New York State Museum Bulletin 321, Albany, NY: University of the State of New York.

Grimm, G.C. 1984: Fire and other factor controlling the big woods vegetation of Minnesota in the mid-nineteenth century. Ecological Monographs 54, 291-311.

He, H.S., Mladenoff, D.J., Sickley, T.A. and 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. and Kucera, C.L. 1956: Composition of pre-settlement forests in three counties of Missouri. Bulletin of the Torrey Botanical Club 83, 207-217.

Hutchison, M. 1988: A guide to understanding, interpreting, and using the public land survey field notes in Illinois. Natural Areas Journal 8(4): 245-255.

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.

Iverson, L.R. and Risser, P.G. 1987: Analyzing long-term changes in vegetation with geographic information system and remotely sensed data. Advances in Space Research 7(11), 183-194.

Jackson, S., Pinto, F., Malcolm, J. and Wilson, E. 2000: A comparison of pre-European settlement (1857) and current (1981-1995) forest composition in central Ontario. Canadian Journal of Forest Research 30, 605-612.

Janke, R.A., McKaig, D. and Raymond, R. 1978: Comparison of presettlement and modern upland boreal forest on Isie Royale National Park. Forest Science 24, 1l5-12l.

Johnston, C.A. 1998: Geographic information system in ecology. Malden, MA: Blackwell Science Inc.

Lam, N. S.-N. and Quattrochi, D.A. 1992: On the issues of scale, resolution, and fractal analysis in the mapping sciences. The Professional Geographer 44, 88-98.

Laurini, R. and Thompson, D. 1992: Fundamentals of spatial information systems. London: Academic Press.

Lee, K.H. and Lunetta, R.S. 1995: Wetlands detection methods. In Lyon, J.G. and McCarthy, J., editors, Wetland and Environmental Applications of GIS, Boca Raton, FL: CRC Press, 249-284.

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

Liegel, C. 1982: The pre-European settlement vegetation of the Aldo Leopold Memorial Reserve. Transactions of the Wisconsin Academy of Science, Arts and Letters 70, 13-26.

Lindsey, A.A. 1961: Vegetation of the drainage-aeration classes of northern Indian soils in 1830. Ecology 42, 432-436.

Loeb, R.E. 1987: Pre-European Settlement Forest Composition in East New Jersey and Southeastern New York. American Midland Naturalist 118, 414-423.

Lorimer, G.G. 1977: The presettlement forest and natural disturbance cycle of northeastern Maine. Ecology 58, 139-148.

Lutz, H.J. 1930: Original forest composition in northwestern Pennsylvania as indicated by early land survey notes. Journal of Forestry 28, 1098-1103.

Maguire, D.J. and Dangermond, J. 1991: The functionality of GIS. In Maguire, D.J., Goodchild, M.F., and Rhind, D.W., editors, Geographical information systems: principles and applications, New York, NY: John Wiley and Sons Inc., 319-335.

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

Manies, K.L., Mladenoff, D.J. and Nordheim, E.V. 2001: Assessing large-scale surveyor variability in the historic forest data of the original U.S. Public Land Survey. Canadian Journal of Forest Research 31, 1719-1730.

Marble, D.F., Calkins, H.W. and Peuquet, D.J. 1984: Basic readings in geographic information systems. Williamsville, NY: SPAD Systems Ltd.

Marks P.L. and Gardescu. S. 1992. Vegetation of the central Finger Lakes region of New York in the 1790s. In Marks, P.L., Gardescu, S. and Seischab, F.K., editors, Late eighteenth century vegetation of central and western New York State on the basis of original land survey records, New York State Museum Bulletin 484, Albany, NY: New York State Museum, University of the State of New York, 1-35.

Marschner, F.J. 1974: The original vegetation of Minnesota (map). St. Paul, MN: USDA Forest Service, North Central Forest Experiment Station.

Martin, D. 1996: Geographic information systems: socioeconomic applications, second edition. London: Routledge.

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.

Mendelson, J. 2002: An additional source of data on northeastern Illinois woodlands around the time of settlement. American Midland Naturalist 147, 279-286.

Mladenoff, D.J., Dahir, S.E., Nordheim, E.V., Schulte, L.A. and Guntenspergen, G.G. 2002: Narrowing historical uncertainty: probabilistic classification of ambiguously identified tree species in historical forest survey data. Ecosystems 5, 539-553.

Nelson, J.C., Redmond, A. and Sparks, R.E. 1994: Impacts of settlement on floodplain vegetation at the Confluence of the Illnios and Mississippi Rivers. Transactions of the Illinois State Academy of Science 87, 117-133.

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

Palik, B.J. and Pregitzer, K.S. 1992: A comparison of presettlement and present-day forests on two Bigtooth aspen-dominated landscapes in Northern Lower Michigan. American Midland Naturalist 127, 327-338.

Pattison, W.D. 1956: Use of the U.S. public land survey plats and notes as descriptive sources. The Professional Geographer 8, 10-14.

Pattison, W.D. 1964: Beginnings of the American rectangular land survey system, 1784-1800. Chicago, IL: University of Chicago Press.

Qi, Y. and Wu, J.G. 1996: Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. Landscape Ecology 11, 39-49.

Radeloff, V.C., Mladenoff, D.J. and Boyce, M.S. 2000: A historical perspective and future outlook on landscape scale restoration in the northwest Wisconsin pine barrens. Restoration Ecology 8, 119-126.

Radeloff, V.C., Mladenoff, D.J., He, H.S. and Boyce, M.S. 1999: Forest landscape change in the northwestern Wisconsin Pine Barrens from pre-European settlement to the present. Canadian Journal of Forest Research 29, 1649-1659.

Radeloff, V.C., Mladenoff, D.J., Manies, K.L. and Boyce, M.S. 1998: Analyzing forest landscape restoration potential: pre-settlement and current distribution of oak in the northwest Wisconsin Pine Barrens. Transactions of the Wisconsin Academy of Sciences, Arts and Letters 86, 189-206.

Russell, E.W.B. 1981: Vegetation of northern New Jersey before European settlement. American Midland Naturalist 105, 1-12.

Schafale, M.P. and Harcombe, P.A. 1983: Presettlement vegetation of Hardin County, Texas. American Midland Naturalist 109, 355-366.

Schneider, M. 1997: Spatial data types for database systems: Finite resolution geometry for geographic information systems. In Goos, G., Hartmanis, J. and Van Leeuwen, J., editors, Lecture Notes in Computer Science 1288, Berlin: Springer.

Schulte, L.A. and Barnes, W.J. 1996: Presettlement vegetation of the lower Chippewa River valley. The Michigan Botanist 35, 29-38.

Schulte, L.A. and Mladenoff, D.J. 2001: The original U.S. public land survey records: their use and limitations in reconstructing pre-European settlement vegetation. Journal of Forestry 99, 5-10.

Schwartz, M.W. 1994: Natural distribution and abundance of forest species and communities in northern Florida. Ecology 75, 687-705.

Sears, P.B. 1925: The natural vegetation of Ohio. I. A map of the virgin forest. Ohio Journal of Science 25, 139-149.

Seischab, F.K. 1990: Presettlement forests of the Phelps and Gorham Purchase in western New York. Bulletin of the Torrey Botanical Club 117, 27-38.

Seischab, F.K. 1992: Forests of the Holland Land Company in western New York, circa 1798. In Marks, P.L., Gardescu, S. and Seischab, F.K., editors, Late eighteenth century vegetation of central and western New York State on the basis of original land survey records, New York State Museum Bulletin 484, Albany, NY: New York State Museum, University of the State of New York, 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.

Seischab, F.K. and Orwig, D. 1991: Catastrophic disturbances in the presettlement forests of western New York. Bulletin of the Torrey Botanical Club 188, 117-122.

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

Shipley, B. and Keddy, P.A. 1987: The individualistic and community-unit concepts as falsifiable hypotheses. Vegetation 69, 47-55.

Shotola, S.J. and Weaver, G.T. 1992: Sugar maple invasion of an old-growth oak-hickory forest in southwestern Illinois. American Midland Naturalist 127: 125-138.

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

Sinton, D.F. 1978: The inherent structure of information as a constraint in analysis. In Dutton, G., editor, Harvard Papers on Geographic Information Systems 6, Reading, MA: Addison-Wesley, 1-17.

Spurr, S.H. 1951: George Washington, surveyor and ecological observer. Ecology 32, 544-549.

Stearns, F. W. 1949: Ninety-years change in a northern hardwood forest in Wisconsin. Ecology 30, 350-358.

Stevens, S.S. 1946: On the theory of scales of measurement. Science 103, 677-680.

Stewart, L.O. 1935: Public land surveys: history, instructions, methods. Ames, IA: Collegiate Press.

Strahler, A.H. 1978: Binary discriminant analysis: A new method for investigating species-environment relationship. Ecology 59, 108-116.

Stritch, L.R. 1990: Landscape-scale restoration of barrens-woodland within the oak-hickory forest mosaic. Restoration and Management Notes 8, 73-77.

Trewartha, G.T. 1932: The Prairie du Chien Terrace: Geography of a Confluence Site. Annals of the Association of American Geographers 22, 119-158.

Truesdell, W.A. 1908: The rectangular system of surveying. Journal of the Association of Engineering Societies 41, 207-230.

Turner, M.G. and Gardner, R.H. 1991: Quantitative methods in landscape ecology. New York: Springer-Verlag.

Van Deelen, T. R. V., Pregitzer, K.S. and Haufler, J.B. 1996: A comparison of presettlement and present-day forests in two northern Michigan deer yards. American Midland Naturalist 135, 181-194.

Veatch, J.T. 1925: Soil maps as a basis for mapping original forest cover. Michigan Academic Science 15, 267-273.

Veregin, H. 1999: Data quality parameters. In Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., editors, Geographical information systems: principles, techniques, management and applications, New York, NY: John Wiley and Sons Inc., 177-189.

White, C.A. 1984: A history of the rectangular survey system. Washington DC: Bureau of Land Management, Government Printing Office.

White, M.A. and 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 presettlement forests of northeastern Ohio. Botanical Gazette 143, 225-237.

Whitney, G.G. 1986: Relation of Michigan's presettlement pine forests to substrate and disturbance history. Ecology 67, 1548-1559.

Whitney, G.G. 1987: An ecological history of the Great Lakes forest of Michigan. Journal of Ecology 75, 667-684.

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

Whitney, G.G. 1996: From coastal wilderness to fruited plain: a history of environmental change in temperate North America from 1500 to the present. London: Cambridge University Press.

Whitney, G.G. and Davis, W.C. 1986: From primitive woods to cultivated woodlots. Thoreau and the forest history of Concord, Massachusetts. Journal of Forest History April, 70-81.

Whitney, G.G. and DeCant, J. 2001: Government land office survey and other early land surveys. In Egan, D. and Howell, E.A., editors, The historical ecology handbook: a restorationist's guide to reference ecosystems, Covelo, CA: Island Press, 147-172.

Whitney, G.G. and Steiger, J.R. 1985: Site-factor determinants of the presettlement prairie-forest border areas of north-central Ohio. Botanical Gazette 146, 421-430.

Wuenscher, J. E. and 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. New Haven, CT: Yale University Press.

Zhang, Q., Pregitzer, K.S. and Reed, D.D. 1999: Catastrophic disturbance in the presettlement forests of the upper peninsula of Michigan. Canadian Journal of Forest Research 29, 106-114.

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

Table 1 Geographic information has the following characteristics and sub-characteristics.

Characteristic	Sub-characteristic
Space	Location
	Geometry (Shape, Size/resolution)
	Topology (Adjacency, Connectivity)
Theme	Nominal
	Ordinal
	Ratio
	Cyclical
Time	Time of observation
	Age of observed object

The final version of the paper has been published in *Progress in Physical Geography* 2005, 29(4): 568-598.
Please contact geowyc@nus.edu.sg for reprint.

Figure 1 The Hardenburgh Patent in Catskill Mountain, New York, an example of the metes and bounds surveys. G.L. is the abbreviation of 'Great Lots' and they were of various sizes and shapes. (Modified from McIntosh, 1962, p.412)

Figure 2 Land division in the GLO surveys.

Figure 3 The HLC survey of western New York, an example of the private land surveys. Ranges are indicated in Roman numerals. Township numbers are to the right. Note its irregularity relative to the GLO surveys.

Figure 4 An example of mislocation of corners. Circles represent the mislocated corners at which the GLO survey posts were found. (Modified from Bourdo, 1956, p.760)

Figure 5 Mislocation of landscape features using a hypothetical example of river course and wetland distribution. (a) The locations of landscape features within a section interpolated by surveyors based on what they encountered along the surveyed section lines. (b) The actual locations of the river and the wetland.