Mapping a forest mosaic

A comparison of vegetation and bird distributions using geographic boundary analysis

Kimberly R. Hall[1] & Susan L. Maruca[2]
[1]*School of Natural Resources & Environment, University of Michigan, Ann Arbor, MI 48109-1115, USA
(e-mail: krhall@umich.edu);* [2]*BioMedware, Inc., 516 North State Street, Ann Arbor, MI 48104-1236, USA
(e-mail: maruca@biomedware.com)*

Abstract

Many areas of ecological inquiry require the ability to detect and characterize change in ecological variables across both space and time. The purpose of this study was to investigate ways in which geographic boundary analysis techniques could be used to characterize the pattern of change over space in plant distributions in a forested wetland mosaic. With vegetation maps created using spatially constrained clustering and difference boundary delineation, we examined similarities between the identified boundaries in plant distributions and the occurrence of six species of songbirds. We found that vegetation boundaries were significantly cohesive, suggesting one or more crisp vegetation transition zones exist in the study site. Smaller, less cohesive boundary areas also provided important information about patterns of treefall gaps and dense patches of understory within the study area. Boundaries for songbird abundance were not cohesive, and bird and vegetation difference boundaries did not show significant overlap. However, bird boundaries did overlap significantly with vegetation cluster boundaries. Vegetation clusters delineated using constrained clustering techniques have the potential to be very useful for stratifying bird abundance data collected in different sections of the study site, which could be used to improve the efficiency of monitoring efforts for rare bird species.

Introduction

Traditionally, plant ecologists have focused their efforts on study sites that are relatively homogeneous with respect to species composition and age. Studies may investigate heterogeneity in plant communities by comparing sampling units that support different plant communities, but typically the goal is to minimize within-site variation so that effects of or differences in plant community can be detected statistically. Although this approach is useful for addressing a variety of research questions, many other areas of ecological inquiry require the ability to detect and characterize change, or boundaries, in ecological variables across both space and time (Turner & Gardner 1991; Wiens et al. 1985; Wiens 1995). Vegetation boundaries can be associated with gradients in one or more environmental variables, or can be caused when natural or human-related disturbances (e.g., fire or logging)

change the distribution of some ecological variable over part of an area (Hansen & diCastri 1992; Johnston et al. 1992; Fortin & Drapeau 1995; Fortin et al. 1996). The ability to create maps of vegetation boundaries has many applications, including use as a baseline for climate or land use change studies, as a source of data for pattern analysis, as a template for studying the relationships among spatial variables, and as habitat maps for animal species. Here we demonstrate how two tools from the field of geographic boundary analysis can be used to characterize and map vegetation patterns in a wetland forest mosaic.

The boundary detection techniques we present here arose primarily from work related to processing and classifying images obtained via remote sensing. Algorithms for image textural analysis were needed to identify areas of rapid spatial change in the images (Hobbs & Mooney 1990; Musick & Grover 1991; Quattrochi

& Pelletier 1991; Fortin 1999). Identifying 'edges' has allowed researchers to extract locations of linear features from the images, such as vegetation boundaries, roads, and geologic contacts. As work in the field of edge detection has progressed, new methods have been developed that allow users to take advantage of multivariate data, including multi-spectral remotely-sensed images (Turner et al. 1991; Johnston et al. 1992; Fortin 1994; Fortin et al., 1996). Here we focus on two of these multivariate methods, multivariate clustering algorithms, and difference boundary delineation, or 'wombling' (Womble 1951; Fortin 1994). Although our application identifies boundaries in field-based vegetation data, the approach is also relevant to the analysis of remotely sensed imagery.

The ability to detect boundaries using more than one vegetation variable can be particularly useful for mapping complex landscapes. In particular, researchers working in areas with highly variable soil conditions, and where disturbance is common may find that using a suite of variables provides a great deal of flexibility in terms of how vegetation is represented in a map. Using boundary detection tools, many different types of vegetation categories can be identified, and the types are not restricted to the observer's predetermined list of classes. In addition, the landscape can be characterized not only by the values of vegetation variables in different areas (e.g., density of a particular overstory species or size class), but also by the variability of a particular variable (or set of variables) across space. Through examining both the location of boundaries identified by various methods, and investigating how various variables contributed to these boundaries, the heterogeneity of landscapes becomes the focus of attention, which we suggest will be very useful in studies of interactions between plant associations in dynamic landscapes.

Geographic boundary analysis typically involves two steps. The first step, boundary detection, identifies spatial locations where a variable or variables show large differences when values at neighboring sites are compared. The second step is statistical evaluation of these boundary features (areas with high rates of change across space) using measures such as boundary contiguity or overlap between two sets of boundaries (Fortin 1994; Fortin et al. 1996; Jacquez et al. 2000). A key difference between boundary analysis and traditional vegetation mapping methods (e.g., those discussed in Küchler 1967) is the focus on the role of space. By explicitly considering space, we can produce maps that highlight within-site spatial variation.

Maps that show and describe heterogeneity are likely to offer many insights into the processes driving observed ecological patterns, and may be particularly useful tools for tracking changes in vegetation over time (Stohlgren et al. 2000).

The two boundary detection methods we demonstrate here, spatially constrained cluster analysis and difference boundary detection (wombling), identify edges in different ways. Spatially constrained cluster analysis is a modification of the more familiar methods of 'ordinary' unconstrained cluster analysis, which group data based on similarity in measured attributes. Constrained clustering groups sampling locations into clusters based on both their geographic proximity and the similarity of their attributes, with boundaries placed between the resulting clusters (Legendre & Legendre 1983; Fortin & Drapeau 1995; Gordon 1999). In effect, these cluster-based approaches produce maps that show bounded areas of relative homogeneity.

In contrast, difference boundary detection first requires estimation of surface gradients (or the equivalent for categorical data), and then identifies locations with the highest rate of change (or slope) as boundary elements. To date, the determination of how large a slope value is required for a location to be part of a boundary is achieved through an arbitrary threshold (usually the top 5–20% of locations), although attempts have been made to find more objective means of selecting boundary locations (Jacquez & Maruca 1998). After selecting boundary elements, decisions on whether to link adjacent boundary elements to form subboundaries (Jacquez et al. 2000; also called subgraphs by Oden et al. 1993) can be made. Connection criteria that evaluate similarity in surface gradient directions are used for this purpose (Oden et al. 1993; Fortin 1994; Fortin et al. 1996). The first criterion prevents two adjacent locations with surface gradients in opposite directions from being linked. The second criterion prevents linkages between two locations for which a link would occur parallel to the direction of change, which is important because a boundary link implies that the change is *across* (i.e., perpendicular to) the link. As with the slope value threshold, connection criteria for linking boundary elements are also typically arbitrary.

Linking boundary elements in this way allows for the assessment of boundary contiguity, or 'cohesiveness', defined as the degree to which boundary elements are linked to form relatively few, highly connected subboundaries. Evaluating cohesiveness is

essential because wombling with an arbitrary threshold always identifies areas with 'large' slope, even if there is no real boundary present. If a true continuous boundary is present, the linked boundary elements should extend over space in a cohesive, non-random pattern (Oden et al. 1993; Fortin 1994). However, if areas with high rates of change are scattered across the landscape, or if adjacent points show different directions of change, boundary elements will not be linked and the 'boundary' will appear as a set of isolated points and/or short line segments. Such would be the case, for example, if the surface were either noisy or mostly flat, or if the scale of measurement was coarser or finer than the phenomenon under study.

The purpose of this study was to investigate ways in which the geographic boundary analysis techniques of spatially agglomerative clustering and difference boundary detection could be applied to create vegetation maps for a forested wetland in northern Michigan. This wetland can be thought of as a mosaic dominated alternately by hardwoods and conifers. The study site also contains a large windthrow area where the density of standing trees is low, but understory vegetation is quite dense. These strong differences in vegetation composition and density over space make vegetation data from this site well-suited to boundary analysis. To illustrate one of the potential uses of these vegetation maps, the distribution of six forest bird species with different habitat preferences are also analyzed using the same tools, and compared with the vegetation maps.

To begin our demonstration, we used spatially constrained clustering to delineate forested areas into ecologically meaningful subsections. Next, we applied the second method, difference boundary detection, with the goal of identifying areas of distinct change over space. With the results of the difference boundary delineation, we conducted a statistical analysis of difference boundary contiguity (or cohesiveness). Cohesive vegetation boundaries suggest that observed patterns in plant distributions are due to some underlying ecological factor, such as a difference in soil type or moisture availability, and are often perceived as ecotones (Johnston et al. 1992; Fortin 1994). If vegetation difference boundaries are not significantly cohesive, the heterogeneity in the forest may be more fine-grained than the scale on which boundaries were sought. Fine-grained patterns of change across space might be found if variation in plant distributions were primarily due to small-scale disturbances, such as treefall gaps (White & Pickett 1985). Of course, both

of these types of boundaries (cohesive and 'broken') could be present at once, as the result of a structurally complex forest. Whether or not a particular analysis technique would detect one or both of these patterns depends primarily on the scale of vegetation sampling (Fortin 1999). One of the major challenges of boundary analysis is that of choosing the appropriate scale for the ecological question addressed; the choice of sampling scale can dramatically alter the results of a geographic boundary analysis (Fortin 1999; Jacquez et al. 2000). These steps completed the boundary detection and analysis component of our work; the next steps focused on examination of relationships between vegetation and bird boundaries using overlap analysis techniques (Jacquez 1995; Fortin et al. 1996).

Many studies have shown that the distribution of breeding forest bird species can be closely linked to a particular combination of tree species or vegetation structure (e.g., Smith 1977; James & Wamer 1982; Freemark & Merriam 1986). To investigate this possibility in our forest mosaic, we compared difference boundaries for bird distribution data with the vegetation boundaries (difference boundaries and cluster boundaries) using boundary overlap analysis (Jacquez 1995; Fortin et al. 1996). Strong overlap would indicate that the differences in vegetation are indeed ecologically meaningful as a factor contributing to the distribution and abundance of breeding birds in this system. Identifying associations between bird species boundaries and vegetation boundaries could be very helpful in the development of monitoring plans and conservation strategies for rare species.

To demonstrate the practical use of the vegetation classification via spatially constrained clustering, we examined the use of these clusters for stratification of bird survey samples. The ability to do post-hoc stratification of the vegetation could help reduce the variance associated with estimates of mean relative abundance (per species) for the total study site (Scheaffer et al. 1990). This is important because, in general, estimates of mean relative bird abundance per point tend to be low (i.e., less than one), while variances are relatively high (Verner 1985; Hall 1996). The purpose of stratifying samples would be to minimize the within-strata variance, and maximize the variance between strata (Scheaffer et al. 1990). If birds associate closely with vegetation clusters, as we expected they should, then the clusters should be useful in stratifying samples of bird data, while also representing a describable habitat unit that can be compared to other studies or other years at the same site. If the stratification is effective,

conservation biologists will be able to calculate means with lower associated variances, which will facilitate detection of changes in relative abundance over time.

Methods

Study area

The study was conducted at the University of Michigan's Biological Station (UMBS), in the northern lower peninsula of Michigan, USA, in Emmett and Cheboygan Counties (45°35′ N lat., 84°42′ W long). We focused on a 45 ha section of the western half of the largest contiguous forested wetland area at UMBS, known as Reese's Swamp. Much of this area was logged and exposed to post-logging fires in the late 1800s, but the forest has regenerated with little human influence since about 1920 (Gates 1942). Most of the wetland area consists of *Thuja occidentalis* L. (nomenclature based on Barnes & Wagner 1981) swamps that are fringed by mixed hardwood uplands, and bisected by numerous small streams. Adjacent to areas of *T. occidentalis* swamp are patches of mature hardwoods; the hardwood areas are often associated with sand ridges that radiate west from Little Carp River, a small river on the east side of the study area. Where the canopy is closed, the understory saplings and shrubs tends to be sparse. Ground cover in these areas consists primarily of leaf litter with sparse ferns, moss, and herbaceous vegetation. In canopy gap areas, understory vegetation is often extremely dense, and is dominated by *Abies balsamea* (L.) Miller in some areas, and *Acer rubrum* L. and other deciduous species in others. Of particular note, there is a large windthrow area in the north-eastern section of the plot with few standing trees and a high density of deciduous understory vegetation. There also are dense patches of *Alnus rugosa* (Du Roi) Sprengel in the wettest areas of the swamp. At the southern boundary of the swamp is Burt Lake; the southern-most sample points are within 100 m of open water. Within this 45 ha area of forested wetland, a transect system was established with permanent markers placed on a 50 m-interval regular grid.

Field methods

Vegetation sampling

We measured forest vegetation following the 0.04 ha circular plot (11.3 m radius) method of James & Shugart (1970) in July and August of 1995. This method for collecting vegetation data was chosen because it is commonly used in studies that assess bird habitat. The study site was divided into 177 square sections (50 m on each side), with permanent location markers placed at the corner of each grid square. Within these squares, vegetation sample points were randomly determined, with the constraint that the entire sample must fit within the grid square. The species and size class of all trees with a dbh ≥ 7.5 cm in the circular plots were recorded, however this analysis uses only the number, and not the size distribution or basal area, for each species. Shrub and sapling density was estimated by counting the number of stems < 7.5 cm dbh of each species along two 22.6 m long by 1 m wide transects (north-south and east-west) through the center (which is counted only once) of the circular plot. James and Shugart's method also included measures of ground cover, canopy cover, and canopy height, but we chose to use only data on overstory (tree) and understory (shrubs and saplings) because we felt they were most likely to be strongly related to bird species' abundances and distributions.

Bird sampling

Point counts, in which all birds heard or seen at a specific location within a given interval of time are recorded (Ralph et al. 1993) were conducted between 0615 and 1130 EDT from June 3–July 19, 1995. Sunrise at UMBS occurred between approximately 0520 and 0550. We conducted 73 point counts, using the permanent location markers within Reese's Swamp to identify parallel transect lines that ranged from 300 to 750 m in length, and were 150 m apart. The first point on each transect was located at a randomly selected distance (between 0 and 40 m) from the start of the transect, and subsequent points were sampled at 150 m intervals. Although no two transects visited on the same sampling day were less than 150 m apart, transects completed in July often crossed or came within 150 m of points sampled in June. All counts lasted for 10 minutes total, and all birds that could be heard or seen were recorded. For this application, only the birds detected within 70 m were tabulated. Because the

point of this study was to examine the spatial boundaries of both birds and vegetation, and vegetation was variable and measured at a small scale (11.3 m radius circular sample on a 50 m grid), this small-scale bird measure appeared appropriate. Reducing the radius of the counts to an even smaller size that was a closer match to the scale of the vegetation measures would lead to smaller sample sizes per point (Hall 1996). In choosing what form of bird data to use, we have attempted to maximize the number of birds observed while minimizing overlap between sampling points and keeping the amount of area sampled similar to the scale of the vegetation data.

We used data for the six most common warbler species in our comparison of overlap between bird boundaries and vegetation clusters and boundaries. We focused on six common species with similar abundance levels so that we could do the analyses without having to standardize or weight species to account for differences in abundance. These six species (Black-and-white Warbler *Mniotilta varia*, Black-throated Green Warbler *Dendroica virens*, Nashville Warbler *Vermivora ruficapilla*, Canada Warbler *Wilsonia canadensis*, Ovenbird *Seiurus aurocapillus*, and Northern Waterthrush *Seirus noveboracensis*) were expected, based on known habitat associations (e.g., as described in Collins et al. 1982; Brewer et al. 1991), to show different patterns of distribution within the study site.

Analysis methods

Data sets

To represent overstory vegetation, we began by treating the number of each of the five most common overstory species recorded in the 0.04 ha samples as separate variables. The choice of how to group the data is important because, unless a specific weighting scheme is chosen, each individual variable receives an equal weight in determining spatial boundaries. These five species, in order of total abundance, were *T. occidentalis*, *Populus tremuloides* Michaux, *Betula papyrifera* Marsh., *Abies balsamea* (L.) Miller, and *Acer rubrum*. Less-common overstory tree species counts were grouped into either 'other deciduous species' [*Fraxinus nigra* Marsh., *Prunus serotina* Ehrhart, *Quercus rubra* L., *Tilia americana* L., *Ulmus americana* L., *Amelanchier arborea* (Micheaux f.) Fernald, *Alnus rugosa*, and *Salix* spp.] or 'other coniferous

species' [including *Pinus strobus* L., *Pinus resinosa* Aiton, *Picea glauca* (Moench) A. Voss, *Picea mariana* (Miller) BSP, and *Larix laricina* (Du Roi) K. Koch]. The understory vegetation was grouped into two variables, deciduous or coniferous. The most common deciduous understory species were *A. rubrum*, *Acer spicatum* Lamarck, *F. nigra*, *A. aborea*, *A. rugosa*, *B. papyrifera*, *P. tremuloides*, *Cornus stolonifera* Michaux, and *Ilex verticillata* (L.) Gray, while coniferous understory stems were mostly *A. balsamea*, with lesser amounts of *T. occidentalis*, *P. glauca*, *P. mariana*, and *P. strobus*. Since the vegetation data were regularly spaced, we were able to treat them as a grid coverage with a pixel size of 50 m by 50 m.

The bird data set had six variables, representing the number of birds detected of each of the six focal warbler species. Because breeding songbirds are territorial (leading to relatively even spacing across the landscape), and territory sizes are generally larger than the 70 m radius circle sample unit, the abundance of birds of the same species at a point tends to be low (ranging from 0–3). For analysis, the bird data were represented as a point coverage, with the center of the 70 m radius area covered in the count treated as the sample location in analyses and figures.

Spatially constrained agglomerative clustering

Forest types and their boundaries were detected using a spatially constrained clustering technique similar to that described by Legendre (1987) and used to define tree communities (Legendre & Fortin 1989; Fortin & Drapeau 1995). In seeking a partition for a given number of clusters, we first used a spatially constrained agglomerative clustering method, using flexible-link linkage (also called intermediate linkage; see Legendre & Legendre 1983 for details) to calculate cluster dissimilarities. This method requires specification of a connectedness parameter, which determines the 'distance of fusion' for a pair of clusters. Distance of fusion is the dissimilarity value used to select the pair of clusters to be merged at each iteration. It is calculated for each pair of adjacent clusters by first enumerating and ordering all pairwise dissimilarities between objects (i.e., sample locations) in one cluster with objects in the other. The distance of fusion is then the (Cn_1n_2)th dissimilarity, where C is the connectedness parameter and n_1 and n_2 are the numbers of objects in the two clusters. We chose a connectedness value of 0.8, which is closer to complete link clustering on the continuum spanned by

the single link (low connectedness) and complete link (high connectedness) methods, because we wanted to emphasize internal cohesion of clusters over external isolation (Gordon 1999). To calculate similarities between pairs of locations we used the Bray and Curtis (also called Steinhaus) dissimilarity measure, as it is self-normalizing and particularly appropriate for frequency data (Legendre & Legendre 1983). After the agglomerative step was complete, we used a k-means clustering algorithm to refine the partition. The k-means method is an iterative procedure that adjusts cluster memberships in the k clusters to minimize the within-cluster sum-of-squares error (SSE). Although spatially constrained k-means cannot guarantee that a global optimum partition will be found, the probability of finding the global optimum (or a partition very similar to it) is increased by starting with a close approximation, such as that achieved through spatially constrained agglomerative clustering.

A common challenge when conducting multivariate clustering is that of deciding how many clusters are actually represented in the data (Milligan & Cooper 1985; Gordon 1999). To aid our selection of number of clusters for the final partition, we found the 'best' partition for cluster counts between 5 and 25 (inclusive), and for each we calculated a goodness-of-fit index $G(k)$ described originally by Caliński & Harabasz (1974) for use in unconstrained clustering applications. Milligan & Cooper (1985) evaluated this index and other stopping rules, and found $G(k)$ to be among the best at correctly identifying the number of clusters in simulated data sets. The index is defined by

$$G(k) = \frac{SSE_B/(k-1)}{SSE_W/(n-k)},$$

where SSE_B is the between-cluster SSE, SSE_W is the within-cluster SSE, n is the total number of objects (in this case, distinct locations) and k is the number of clusters. Local maxima for $G(k)$ identified potential final partitions at 9, 14, 16, 19, and 21 clusters (Figure 1). From these choices we selected 14 as the cluster count for the final partition, for several reasons. Because we were interested in relationships between vegetation types and the spatial distributions of breeding warblers, we wanted the scale of the partition to be no finer than that of warbler territories for the species examined, which are typically no smaller than 0.5 ha. Therefore, we wanted to reject any partition with too many singleton clusters. Because we were also interested in using the partition for stratified sampling, we wanted a relatively low upper limit on the number of
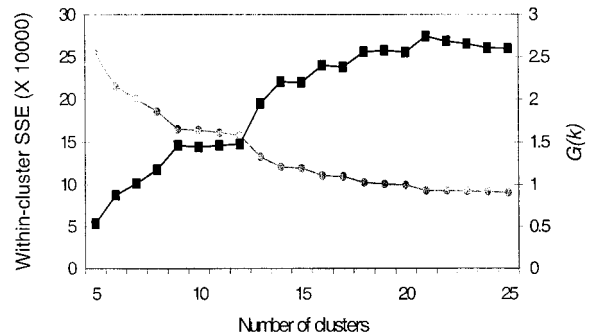

Figure 1. Within-cluster sum of squares error (•) and goodness-of-fit index $G(k)$ (■) for partitions with cluster counts ranging from 5-25. Local maxima in $G(k)$ occurred at 9, 14, 16, and 21 clusters; these partitions were examined visually to obtain a final partition with 14 clusters.

'strata', so that sample sizes within each would be adequate for this application. Partitions with 16, 19, and 21 clusters were deemed too fine-scaled, with too many small clusters, to be appropriate for this study. The partition with 9 clusters did not reflect some of the spatial pattern that we judged could be important to warblers in the selection of territory location.

Difference boundary detection

We used the wombling method for grid (Womble 1951) and irregular point data (triangulation wombling, Fortin 1994) to detect spatial zones of rapid change in vegetation densities and warbler densities, respectively. Throughout the text, we will refer to these zones as difference boundaries (Jacquez et al. 2000). To 'womble' with multiple variables, surface gradients were first estimated for each variable, the magnitudes of which were averaged at each potential boundary location. To distinguish those locations of large average slope (boundary elements), we applied a threshold to each data set that was selected by first examining the distribution of average gradient magnitudes, in the form of a histogram, and looking for natural breaks. This resulted in the designation of 15% of the potential boundary locations as boundary elements for the vegetation data, and 17% for the bird data. Adjacent boundary elements were then connected to form continuous boundaries (or subboundaries) according to two connection criteria: adjacent gradient vectors could differ in direction by no more than 100 deg, and each vector could be no closer in direction to its connecting line than 30 deg.

To evaluate the cohesiveness of the vegetation and warbler boundaries, subboundary statistics were

calculated and evaluated using Monte Carlo procedures. The statistics, developed by Oden et al. (1993), are: number of subboundaries, number of singletons (subboundaries comprised of a single boundary element), maximum and mean length of subboundaries (i.e., number of connected boundary elements), and maximum and mean diameter. Subboundary diameter is defined as the minimum number of connections between the most separated pair of boundary elements within a subboundary. Continuous boundaries delineating pronounced landscape features are characterized by a small number of subboundaries, few singletons, and high mean and maximum lengths and diameters (Fortin & Drapeau 1995). The null model used to evaluate these statistics was that of complete spatial randomness; 500 simulations were performed to amass the null distribution. For these and all other statistical tests, alpha was set at the 0.05 level of significance.

Overlap analysis and sampling example

To examine the coincidence between (1) the warbler difference boundaries and vegetation difference boundaries, and (2) the warbler difference boundaries and vegetation cluster boundaries, we used the distance-based boundary overlap statistics developed by Jacquez (1995). For our work the statistics are:

- O_W: average geographic distance from a boundary element in the warbler boundaries to the nearest boundary element (or cluster edge) in the vegetation boundaries.
- O_V: average geographic distance from a boundary element (or cluster edge) in the vegetation boundaries to the nearest boundary element in the warbler boundaries.
- O_{WV}: average geographic distance from a boundary element in either set of boundaries to the nearest boundary element in the other set.

O_W and O_V measure directional association between boundary sets, while O_{WV} reflects the overall degree of overlap, regardless of direction. Because there exists a large body of ecological evidence to suggest that vegetation directly influences bird distributions (e.g., Smith 1977; Collins et al. 1982; James & Wamer 1982; Cody 1985; Freemark & Merriam 1986), we chose a null model that preserved vegetation boundaries and randomized only the warbler distribution data during Monte Carlo simulations. Again, we used 500 Monte Carlo runs and randomized the original

data to simulate complete spatial randomness of birds recorded at the sampling points. All boundary detection and analysis procedures were performed using the pre-release version of TerraSeer, Inc.'s Boundary-Seer software for geographic boundary analysis (see http://www.terraseer.com for more information).

To demonstrate the potential use of cluster analysis as a tool for stratifying bird abundance data collected on the research site, we chose to reduce the total number of clusters from 14 to 4, a number more practical for this type of task. To do this, we first merged the two island singleton clusters (a one pixel cluster surrounded by a single other cluster) with their respective surrounding clusters. To merge the other singleton, we examined the dissimilarities between each singleton and its neighbors using the Bray and Curtis measure comparing means, and then merged the singleton with its most similar neighbor. Other clusters with low dissimilarities were then grouped together, and final choices in terms of groupings were made based on the examination of means of various vegetation variables, with a focus on grouping clusters with similar understory vegetation densities. To determine mean bird abundance values stratified by cluster, each point count location was assigned to the overlapping vegetation cluster. If the count location was on a cluster boundary, two rules were followed:

- if the count was on the edge of a large cluster group and a smaller group, it was assigned to the smaller group, as these groups had smaller sample sizes;
- if the count was on the edge of two smaller clusters, a coin flip was used to assign it to a vegetation strata.

Three point count locations that were outside of the vegetation sampling area were not used in this part of the analysis.

Results

The distribution of the nine vegetation variables in the study area (Figure 2) demonstrates the heterogeneity of the site vegetation. We determined that the most appropriate partition of the data produced 14 clusters, shown (along with the warbler difference boundaries) in Figure 3. To help in interpretation of the vegetation cluster delineation, the mean and standard error for each vegetation variable were calculated by cluster (Table 1). Through examination of the vegetation values for each cluster (Table 1, Figure 2), it is possible to the identify the vegetation variables driving the bound-
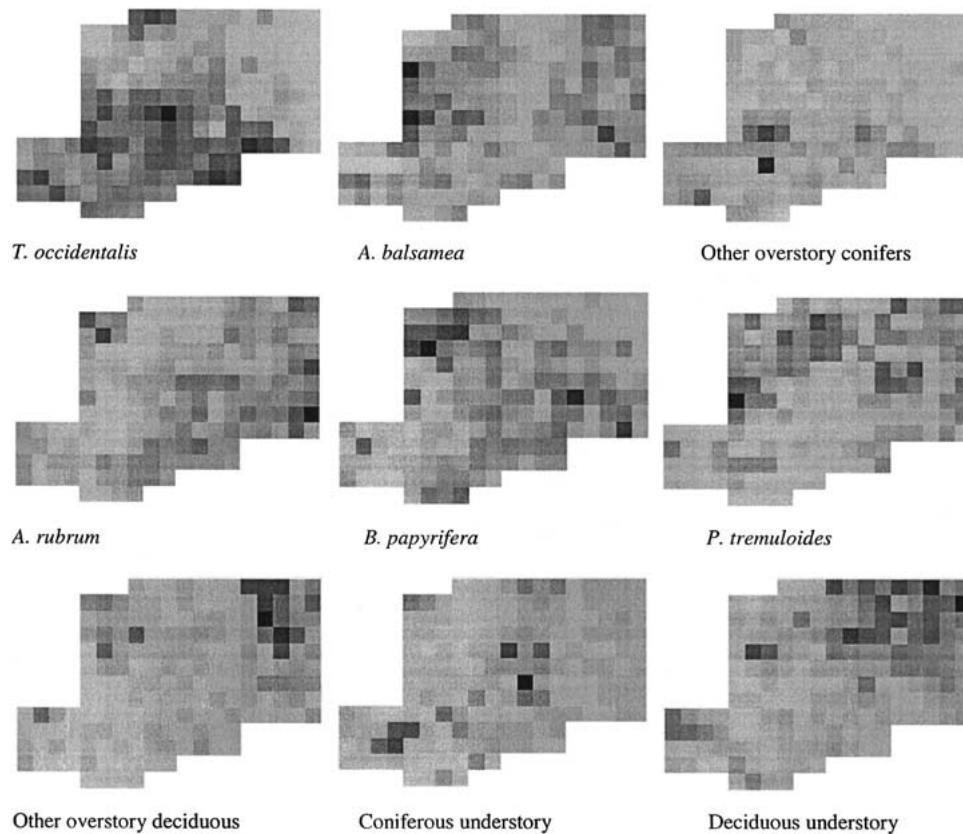
Figure 2. Spatial pattern of each of the nine vegetation variables at the Reese's Swamp site (each grid cell is 50×50 m). Dark values indicate high densities for that variable, while lighter values indicate lower densities. The southern boundary of the site is particularly irregular because it follows the edge of a small lake.

ary placement. For example, the largest cluster (number 14, $N = 89$) was dominated by *T. occidentalis* in the overstory and contains very little understory vegetation. Clusters 2 and 4 are small areas that deviate from the surrounding cluster (10) in that they contain much less deciduous understory. Clusters 6, 7, 8, and 9 collectively represent a highly variable transition zone between clusters 14 (high *T. occidentalis*) and 10 (mostly characterized by high values for deciduous understory). Cluster 11 is similar to 10, except it contains considerably more *T. occidentalis*. Clearly, small areas were commonly identified as clusters – for example, Cluster 12 is dominated by a patch of spruce (high 'other conifers') amongst the *T. occidentalis* (cluster 14).

Using the Bray and Curtis dissimilarity between means for each cluster pair, we determined that the following groups of clusters represented the same vegetation types: clusters 3, 5, and 10, characterized primarily by high counts of deciduous under-

story; clusters 7 and 13, characterized by moderately high *T. occidentalis* densities and a dense coniferous understory; and clusters 1 and 9, characterized by mixed overstory species composition and relatively low understory counts. Characteristics of these cluster groups, which formed the basis of our stratified sampling scheme, are presented in Table 2.

The difference boundaries (with subboundary connections) for vegetation and warblers are shown in Figure 4. The subboundary statistics for the vegetation data (based on a 15% threshold) suggest that the number of subboundaries is significantly smaller than expected under a null hypothesis of complete spatial randomness ($p = 0.03$), while the mean length is significantly longer than expected ($p = 0.03$, Table 3). The mean diameter was nearly significant ($p = 0.056$), which, although far from conclusive by itself, can aid in the interpretation of the statistics as a set. The number of singleton subboundaries, maximum length, and maximum diameter were not

Table 1. Mean (S.E.) number of overstory trees and understory stems at vegetation sample sites in the 14 different clusters, listed by cluster number (see Figure 3). Total area for overstory samples is 0.04 hectares, and total area for understory samples is 44 m^2. N = number of vegetation samples included each cluster.

Cluster	*Thuja occidentalis*	*Abies balsamea*	other conifers	*Acer rubrum*	*Betula papyrifera*	*Populus tremuloides*	other deciduous	Coniferous Understory	Deciduous understory
1	4.4	9.4	1.6	3.2	7.1	4.9	4.7	8.3	8.6
N = 14	(1.0)	(2.8)	(0.59)	(1.2)	(1.3)	(0.85)	(1.0)	(4.4)	(2.0)
2	0	8.5	0	1.5	0	0	12.5	0.5	16.5
N = 2	(0)	(5.5)	0	(0.5)	(0)	(0)	(10.5)	(0.5)	(15.5)
3	0	15	1	0	1	6	2	0	55
N = 1	–	–	–	–	–	–	–	–	–
4	15	1	0	3	2	0	1	0	10
N = 1	–	–	–	–	–	–	–	–	–
5	1	10	1	1	4.5	3.5	8.5	1	94.5
N = 2	(1)	(3.0)	(0)	(0)	(3.5)	(0.5)	(5.5)	(1)	(20.5)
6	1.6	8.4	0.4	3	2.8	6.8	0.4	0.8	30.8
N = 5	(1.4)	(2.3)	(0.24)	(1.2)	(0.73)	(1.9)	(0.24)	(0.37)	(7.0)
7	20.5	1.6	5.5	2.5	1.6	0.38	0.5	53.3	24.6
N = 8	3.5	(1.1)	(2.9)	(1.1)	(0.68)	(0.26)	(0.27)	(14.5)	(7.0)
8	0.5	17.5	0.25	3.5	7.5	0.25	4.25	0.5	4.5
N = 4	(0.29)	(4.7)	(0.25)	(0.96)	(2.6)	(0.25)	(1.6)	(0.29)	(2.6)
9	0.67	10.4	0.11	8.7	3.1	4.11	4.11	4.6	18.2
N = 9	(0.55)	(1.8)	(0.11)	(1.8)	(1.3)	(0.86)	(1.4)	(0.90)	(3.2)
10	3	4.8	0.3	2.2	0.61	3.15	6.97	6.1	70.2
N = 33	(0.97)	(1.0)	(0.10)	(0.44)	(0.23)	(0.60)	(1.4)	(1.5)	(5.1)
11	13.8	0.6	2.6	1.4	2.4	1.2	2.2	5	45.2
N = 5	(4.2)	(0.60)	(1.0)	(0.51)	(1.5)	(1.2)	(2.2)	(1.9)	(6.6)
12	20	6	56	0	0	1	0	0	0
N = 1	–	–	–	–	–	–	–	–	–
13	18.8	4.8	3.5	1.3	0.25	1.8	1.25	77.5	8.5
N = 4	(3.7)	(0.48)	(0.87)	(0.48)	(0.25)	(1.0)	(1.0)	(6.6)	(3.9)
14	41.0	7.4	6.1	2.1	2.4	2.06	0.57	4.5	4.1
N = 89	(2.0)	(0.84)	(0.81)	(0.22)	(0.24)	(0.39)	(0.39)	(1.0)	(0.78)

Table 2. Mean (S.E.) number of overstory trees and understory stems at vegetation sample sites in 3 repeated vegetation 'types', or clusters that showed strong similarities. Total area for overstory samples is 0.04 hectares, and total area for understory samples is 44 m^2. N = number of vegetation samples included each cluster.

Clusters	*Thuja occidentalis*	*Abies balsamea*	other conifers	*Acer rubrum*	*Betula papyrifera*	*Populus tremuloides*	other deciduous	coniferous understory	deciduous understory
3,5,10	2.8	5.5	0.4	2.1	0.8	3.3	6.9	5.6	71.2
N = 36	(0.9)	(1.0)	(0.1)	(0.4)	(0.3)	(0.5)	(1.3)	(1.4)	(4.9)
7,13	19.9	2.7	4.8	2.1	1.2	0.8	0.8	61.3	19.3
N = 12	(2.5)	(1.2)	(1.9)	(0.7)	(0.5)	(0.4)	(0.3)	(10.2)	(5.3)
1,9	3.0	9.8	1	5.3	5.6	4.6	4.5	6.8	12.3
N = 23	(0.8)	(1.8)	(0.4)	(1.1)	(1.0)	(0.6)	(0.8)	(2.7)	(2.0)

Table 3. Subboundary statistics for the vegetation and warbler species boundaries. Significance values are one-tailed in the direction of boundary cohesiveness. Lengths and diameters are expressed in terms of the number of connected boundary elements.

	Number of subboundaries	Number of single BEs	Mean length	Maximum length	Mean diameter	Maximum diameter
	10	6	2.2	7	2.0	6
Vegetation	$p = 0.032$	$p = 0.12$	$p = 0.032$	$p = 0.58$	$p = 0.056$	$p = 0.07$
	9	3	2.3	6	2.2	5
Warblers	$p = 0.22$	$p = 0.38$	$p = 0.22$	$p = 0.16$	$p = 0.21$	$p = 0.32$
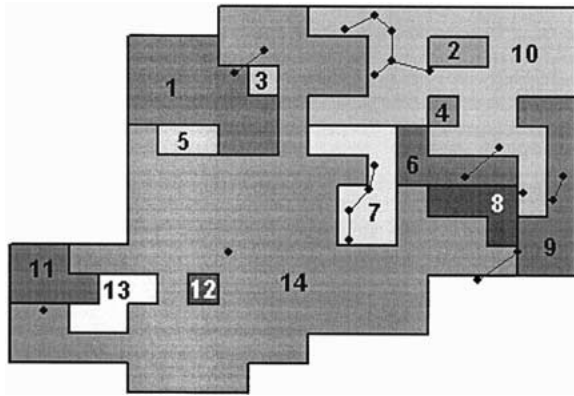
Figure 3. Vegetation clusters obtained from spatially constrained clustering (shown with warbler abundance difference boundaries, see Figure 4). The following combinations of clusters represent the same vegetation type: clusters 3, 5, and 10 (dense deciduous understory, little overstory), clusters 7 and 13 (moderately dense *T. occidentalis* and dense coniferous understory), and clusters 1 and 9 (mixed overstory and relatively low density deciduous understory). See text and tables for other cluster characterizations.
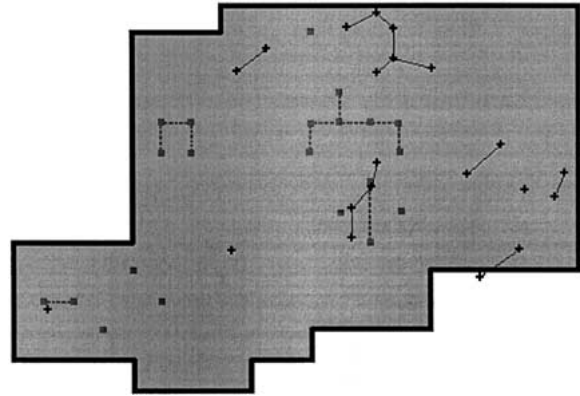

Figure 4. Difference boundaries for vegetation (gray squares, linked with a dashed line) and warbler abundances (black 'x's, linked with a solid line). Boundaries were detected using a 15% threshold for vegetation and 17% for warbler abundances. Lines represent subboundary connections; see text for details regarding connection criteria.

different from those based on a random arrangement of the data. Overall, we concluded that the vegetation boundaries were moderately but significantly continuous. Difference boundaries in the densities of the six warbler species (based on a 17% threshold) were not significantly cohesive according to any of the six subboundary statistics evaluated (Table 3).

We found no significant boundary overlap between the warbler difference boundaries and vegetation difference boundaries (Table 4), suggesting there is no strong spatial association between the areas where bird abundances and tree and shrub abundances change rapidly through space. However, there was significant overlap between warbler difference boundaries and vegetation cluster boundaries (Table 4). The average minimum distance from a warbler boundary element to a vegetation cluster boundary (O_W) was 16 m ($p = 0.03$), and the average distance from a boundary ele-

ment in either set to its nearest neighbor in the other set (O_{WV}) was 19 m ($p = 0.01$). This suggests that the warblers, in their selection of territories, are responding to differences in vegetation communities, but the areas where bird species turnover occurs are not correlated with the areas with the most dramatic changes in vegetation. Not surprisingly, the average minimum distance from a vegetation cluster boundary to a warbler boundary element was not significant, suggesting that the vegetation boundaries were not 'influenced' by the location of the warbler boundaries ($O_V = 23$ m, $p = 0.09$).

The grouped clusters used for stratifying the bird abundance observations varied primarily in terms of the abundance of *T. occidentalis*, and the density of understory vegetation (Table 2, Table 5). The first group, including clusters 1, 8 and 9, consisted of a mixed overstory with few *T. occidentalis* and a medium to high density, mostly hardwood, understory. Northern

Table 4. Overlap statistics. Significance values are one-tailed in the direction of boundary overlap. Distances are expressed in meters.

	O_W	O_V	O_{WV}
Warbler (difference) and	118	86	101
vegetation (difference)	$p = 0.79$	$p = 0.49$	$p = 0.67$
Warbler (difference) and	16	23	19
vegetation (cluster)	$p = 0.03$	$p = 0.09$	$p = 0.01$

Waterthrushes were the only species that had their highest abundances in samples collected within this group (Table 5). Group 2 (clusters 2, 3, 4, 5, 6, 10 and 11) also had few *T. occidentalis*, and in general had the lowest overstory density and highest understory density, especially of deciduous understory understory (singleton cluster # 3 is a high density *A. rugosa* clump). Nashville Warblers were found in the highest densities in group 2, and overall warbler abundance tended to be highest in these areas (Table 5). Group 3 (clusters 7 and 13) and Group 4 (clusters 14 and 12) were both dominated by *T. occidentalis* in the overstory, but group 3 had a dense understory layer dominated by *A. balsamea*, while group 4 had little understory (cluster 12 was isolated from 14 because it is dominated by *P. glauca*, one of the 'other conifers' in the overstory). Three species were most abundant in group 3 (Ovenbird, and Black-and-white and Canada Warblers), while only the Black-throated Green Warbler was most common in group 4 sites.

Discussion

Representing the wetland forest as vegetation clusters

The heterogeneous nature of the vegetation in Reese's swamp site is clear from the spatial patterns of the different variables shown in Figure 1. In general, *T. occidentalis* dominates the southern portion of the site, which is adjacent to a small lake. The areas to the north and east, which are typically drier, tend to be dominated by hardwoods, although there are scattered hardwoods throughout the area. In the past 30–40 years, the relative dominance of conifers in the overstory at Reese's Swamp has decreased (Parody 1996), and numerous small and large treefall gaps have opened, and over time, have been filled with new vegetation. Characterizing the pattern of vegetation in this dynamic, heterogeneous ecosystem is likely to be a challenge with any analysis method, but we suggest that boundary analysis techniques are uniquely suited to help researchers meet this challenge.

Spatially constrained clustering provides a way to treat a vegetation mosaic as a collection of internally homogeneous areas. As described in the methods, the first difficulty encountered in identifying these homogeneous areas often is deciding how many clusters are actually represented in the data (Milligan & Cooper 1985; Gordon 1999). The goodness of fit approach employed here was very helpful in reducing the arbitrariness of this decision, although in our experience with this method we have found that a larger number of clusters than expected was often most useful, as many unexpected single or two-unit clusters may occur. Because we used nine different unstandardized vegetation variables and each was given equal weight, cluster determinations were often driven by the variables with the highest maximum values, i.e., *T. occidentalis*, and deciduous and coniferous understory (Figure 1, Table 1). In some cases, one vegetation sample was so unusual that it led to that particular grid square becoming isolated from all others. For example, cluster 12 was determined by the high density of 'other overstory conifers' (see Figure 1 and Figure 3). Similarly, the location of cluster 5 is driven primarily by a clump of *A. rugosa* in the understory. The strong effect of the two understory variables in the location of boundaries seems appropriate for our application of the vegetation map, since low vegetation is often an important habitat variable influencing the distribution of warblers at the local scale (e.g., Collins et al. 1982; Cody 1985). For other applications, such as a focus on identification of treefall gaps, less emphasis on understory variables might be more appropriate (see below).

Use of clusters identified by these methods requires that additional judgements be made in terms of minimum acceptable cluster size, and maximum cluster dissimilarity for deciding whether two clusters represent the same vegetation type. For the application of stratifying sampling for bird surveys, we chose to combine small clusters into nearby larger ones (e.g., clusters 3, 4, and 12), even when they were somewhat different, since these small areas are typically smaller than an average warbler's territory size and could not be meaningfully sampled. However, in other applications, such as identification of rare plant assemblages for conservation, it might not be desirable to merge small, unique clusters with larger ones. We used a similarity index approach to help us group non-adjacent but similar clusters into

Table 5. Mean bird detections per 10-minute, 70 m radius point count for 6 warbler species observed in the grouped vegetation clusters. N = number of points counts conducted in that cluster. Highest detection rate for each warbler species among the four clusters is in bold.

Species	Combined vegetation clusters			
	Group 1 (1, 8 and 9) $N = 10$	Group 2 (2, 3, 4, 5, 6, 10 and 11) $N = 22$	Group 3 (7 and 13) $N = 6$	Group 4 (12 and 14) $N = 18$
Number of species per point (max = 6)	1.3 (0.47)	2.23 (0.27)	**2.33 (0.67)**	1.14 (0.18)
Total number of warblers per point	2.4 (0.69)	**2.90 (0.38)**	2.83 (0.79)	1.22 (0.20)
Black-and-white Warbler	0.5 (0.17)	0.68 (0.15)	**1.0 (0.44)**	0.31 (0.090)
Black-throated Green Warbler	0.10 (0.10)	0.09 (0.063)	0	**0.26 (0.095)**
Nashville Warbler	0.10 (0.10)	**0.63 (0.12)**	0.5 (0.34)	0.31 (0.080)
Canada Warbler	0.30 (0.15)	0.59 (0.17)	**0.67 (0.21)**	0.11 (0.055)
Ovenbird	0.30 (0.15)	0.18 (0.10)	**0.33 (0.21)**	0.17 (0.065)
Northern Waterthrush	**1.1 (0.43)**	0.73 (0.20)	0.33 (0.21)	0.057 (0.040)

'vegetation types.' This method of simplifying cluster analysis results is likely to be very useful in sites where similar vegetation types show a repeated pattern across the landscape.

Delineating and analyzing difference boundaries

As was true in the cluster results, many of the difference vegetation boundaries identified in our analyses reflected the presence of isolated patches of dense, distinct vegetation. An examination of the vegetation subboundary statistics suggested that these boundaries were significantly but perhaps moderately cohesive; two of six of the diagnostic measures were significant. However, these boundaries do not appear to have been generated by a single ecological factor, such as a change in hydrology or soils, as they were scattered around the study site, and tended to be quite narrow (one or two boundary elements wide). We suggest that these boundaries represent a combination of the transition from a mostly coniferous to mostly deciduous (overstory) swamp (the subboundaries in the center of the study area), and also indicate the pattern of disturbance at the site.

The main type of disturbance influencing vegetation in Reese's Swamp (in the absence of fire and logging) has been treefall gaps, although there were a few small roads and trails. A very large area (about 4–5 ha) that was found in cluster 10 near the zigzagging subboundary in the north-eastern corner had very few standing trees and many trees on the ground, presumably due to a strong wind event. In the upper section of the south-western corner (included in clus-

ter 11) there was another disturbed area that was the site of an old road, and featured a dense deciduous understory. Many of the single boundary elements were likely to be smaller treefall gaps, which, like the larger windthrow areas, can result in high rates of change in both the overstory variables (low relative to surrounding areas) and understory vegetation variables (high relative to surrounding areas, especially when clumps of *A. rugosa* or *A. balsamea* were present). To explicitly map patterns of treefall gaps, which were typically fairly small, we would have used finer resolution data. Similarly, small (5–10 m^2), dense patches of *A. balsamea* saplings often occured in the understory, especially in areas dominated by a fairly dense *T. occidentalis* overstory. Any pattern in these patches would be too fine-grained for this sampling scheme to detect, yet these patterns might provide useful insights into ecological processes that affect both the vegetation and, as a result, the distribution of birds.

In contrast, these samples may have been too fine-grained for detecting the broader-scale transition between vegetation types in the swamp using difference boundary detection. We felt that our sampling scheme (11.3 m radius circles within the 50 m grid squares) would be adequate for detecting the conifer-to-hardwoods transition that we had observed in visits to the site, but perhaps complete inventories would have reduced the impact of small, dense clumps of vegetation. This broad-scale pattern of transition is apparent in the clustering results, which are likely better at picking up broad patterns due to the simple fact that data are combined into larger units through the

clustering process. In our analyses, some isolated areas with very dense clumps of one or more species tended to be detected as boundaries, while more gradual changes in relative dominance were missed, but could potentially be captured if data from individual grid cells were combined into blocks of 4 or more. As suggested by Fortin (1999), these results suggest that a sensitivity analysis focusing on the effect of spatial resolution should be included in future work involving boundary detection methods.

Scale issues notwithstanding, it is important to realize that crisp boundaries are not always present. In many vegetated areas, distinct boundaries may not exist where topography and soils are similar; in others, sharp boundaries between vegetation are typical (Beals 1969). In Reese's Swamp, we are likely to have both sharp and gradual boundaries, the sharp boundaries being better represented by the approaches demonstrated here. When we examined partitions for different total numbers of clusters, we observed that some areas were not consistently classified, while other clusters were persistent among partitions. In particular, the area occupied by clusters 7 and 10 in Figure 3 showed high variation, suggesting that the landscape here is transitional and changes across the landscape are gradual. As a result, no truly distinct pattern emerges in this area. Because one of the dominant elements of this study site is a fairly wide transition zone from a *T. occidentalis*-dominated conifer swamp to hardwood swamp, a fuzzy classification approach (e.g., Burrough 1989; Brown 1998), which allows locations to have partial membership in more than one cluster (or class), would be a useful extension of this research.

There was little evidence for cohesive boundaries in the bird data; none of the diagnostic subboundary statistics was significant. The bird boundaries appear to be located near transitions between forest types (as indicated by the clusters), but not necessarily at the areas with highest rates of change along the transitional areas. This idea was supported by the overlap analysis results, in which the overlap of the bird boundaries with the vegetation difference boundaries was not significant, but overlap of the bird boundaries with the cluster boundaries was significant. Our results suggest that the birds may be avoiding the transition zones when performing their territorial defense (i.e., singing) behaviors, or that they could be responding to more gradual changes in habitat that we did not detect using these methods and this scale of vegetation data.

Applications to management and conservation

Based on our examination of the potential application of boundary analysis as an aid in stratified sampling, this approach could be quite useful. For example, the Northern Waterthrush, a species that would have very low means if all values were pooled, showed a relatively high mean of 1.1 birds per point in points sampled within vegetation group 2 (Table 4). Higher means and lower variance relative to the mean increases the statistical power of tests for differences between means, so these types of stratification schemes could be very helpful for researchers monitoring population changes over time (Cohen 1988; Scheaffer et al. 1990; Hall 1996). Again, it is likely that the importance of understory vegetation in the cluster classification is key to the success of this stratification scheme. As might be predicted from the results in Table 5, Northern Waterthrush and Canada, Nashville, and Black-and-white Warblers all tend to favor areas with dense understory, while Ovenbirds and Black-throated Green warblers are more often found in forests with open understory (Brewer et al. 1991).

Differences in bird abundance means among samples collected from different vegetation clusters are likely to represent both actual differences in abundance, as well as differences in bird detectability. Besides habitat structure, many other factors such as time of day, wind conditions, and time within the breeding season influence the number of birds detected at a point count (e.g., papers in Ralph & Scott 1981). Although grouping bird data into sets with similar detection probabilities (based on boundaries in vegetation) should reduce the variance of bird abundance estimates, applying boundary analysis directly to the bird data is likely to produce less meaningful boundaries due to the challenges associated with accurately counting birds.

Although our example application applies specifically to wildlife management and conservation, boundary analysis methods are likely to be useful in many other applications. Many land managers are faced with the challenges of classifying and managing heterogeneous land areas, and changes in plant distributions are likely to be important to researchers looking for evidence of processes such as climate change (e.g., Millington & Alexander 2000), and other types of natural and human-influenced drivers of vegetation change (e.g., Camarero et al. 2000). A recent paper by Stohlgren et al. (2000) highlights the fact that although many landscape ecologists have called attention to the

need for a greater understanding of the nature of eco-tones and other heterogeneous areas, little progress has been made since this point was emphasized years ago by Wiens et al. (1985). Geographic boundary analysis is likely to be useful in these endeavors, although standard methods for identifying boundaries have yet to be developed. In particular, quantifying heterogeneity by examining patterns of high rates of spatial change could facilitate comparisons of plant distribution patterns among sites or for the same site over time (e.g., Stohlgren et al. 2000; Parody et al., in press).

For researchers creating vegetation maps for the purpose of modeling such factors as land use change, or vertebrate population distribution, flexibility in the number of plant associations identified using boundary analysis could facilitate incorporation of a chosen level of habitat variability into the models. For some applications, this flexibility may render boundary techniques more appropriate for vegetation mapping than the frequently used patch-based methods, which treat various landscapes as mosaics consisting of patch and non-patch (e.g., McGarigal & McComb 1995; Chapin et al. 1998).

Key areas for future research

For multivariate boundary analysis to be truly useful in ecological work, several factors require further examination. One significant challenge involves the development of guidelines for what types of data should be included in a particular type of analysis. The choice of variables will be a key consideration for any attempt to map vegetation using geographic boundary analysis (Fortin 1997, 1999). In this work, we focused on overstory tree densities, with shrub densities also included. We believe that inclusion of the two understory-related variables is likely the main factor contributing to significant overlap of bird boundaries with vegetation clusters. In classifications for other purposes, other variable choices may be more appropriate. Working in a woodland in New York, USA, Fortin (1997) found significant overlap between boundaries detected from different data types (e.g., tree and shrub density, percent coverage, or presence-absence), but also found a significant spatial lag between boundaries based on shrubs alone and trees alone. Fortin's (1997) work highlights how different forms of data can be used to identify different types of boundaries, and how these different boundaries are likely to be linked to different causal mechanisms. She suggests that boundaries based on presence/absence may indicate where species

interactions and physical tolerances are most important, while boundaries based on changes in abundance are more likely due to species responses to variation in environmental quality. In a non-boundary analysis example, work by Zogg & Barnes (1995) on classifying wetland vegetation at UMBS suggests that wetland types are poorly separated on the basis of overstory vegetation, in part due to the effects of human disturbance (selective logging and burning). This work suggests that a more comprehensive approach utilizing ground-flora, hydrology, and soils might provide more robust clusters. Whether this approach would be as appropriate when the goal of characterization is to classify wildlife habitat is worthy of further study.

Similarly, the number of variables included is an important consideration. If a researcher decides to use many, equally-weighted, species variables in a vegetation data set, changes in rare species may be overemphasized relative to their importance in the forest community, while changes in common species may be underemphasized. In our work, we dealt with this problem by limiting the number of our vegetation variables to 9, but this meant that rare species were lumped into the 'other' categories, and any patterns in those species were not likely to strongly influence the boundary determination process. The choice of standardization methods or similarity/dissimilarity measures also influences how variables are weighted. We feel that our selection of variables and our decision to use unstandardized count data with the Bray & Curtis dissimilarity measure were appropriate for this work, and that we were unlikely to overemphasize rare species. However, access to other vegetation data and the use of alternative standardizations and/or dissimilarity measures might uncover slight differences in the detected pattern that might be meaningful for predicting the presence of breeding warblers, or for understanding some aspect of vegetation dynamics in this ecosystem.

In addition to choosing variables and deciding how they are standardized and weighted, the researcher must also focus on how best to implement boundary analysis methods, which involves making choices regarding many different methodological criteria (see introduction and methods). Guidance for making these choices should be a major focus of future work. Specific priorities include: replacing subjective thresholds by more objective criteria during wombling (e.g., see preliminary work by Jacquez & Maruca 1998); developing multiscale approaches to boundary detection that will allow for the characterization of boundary

strength, length, sinuosity, scale, and fuzziness (see the promising progress made here by Csillag et al. 2001); and establishing guidelines for the appropriate use of the various boundary detection techniques, including the advantages, disadvantages, and general interpretation of difference (open) boundaries versus cluster (closed) boundaries.

Conclusions

Maps of vegetation are crucial components of many different types of conservation, modeling, and management efforts. Many new sources of data (e.g., satellite imagery with better resolution), and new data management and visualization tools (e.g., geographic information systems) are becoming widely used, and new tools are needed for ecological questions that deal with heterogeneity over space and time (Millington & Alexander 2000). The tools of boundary analysis and overlap statistics are likely to be useful in many ecological applications that require the interpretation of either remotely-sensed or field-collected spatial data sets. Boundary detection techniques have been identified as a key issue in vegetation mapping (Millington & Alexander 2000), and an area of 'research needed to further conservation' in a review of spatial modeling applications in bird conservation (Villard et al. 1998, p. 58). We hope that this example application of boundary analysis techniques encourages plant ecologists to explore future applications and improvements on these methods in their own work.

Acknowledgements

References

Barnes, B. V. & Wagner, W. H., Jr. 1981. Michigan trees: A guide to the trees of Michigan and the Great Lakes Region. The University of Michigan Press, Ann Arbor, MI.

Beals, E. W. 1969. Vegetational changes along altitudinal gradients. Science 165: 981–985.

Brewer, R., McPeek, G. A. & Adams, R. J., Jr. 1991. The atlas of breeding birds of Michigan. Michigan State University Press, East Lansing, MI.

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

Burrough, P. A. 1989. Fuzzy mathematical models for soil survey and land evaluation. J. Soil Sci. 40: 477–492.

Caliński, T. & Harabasz, J. 1974. A dendrite method for cluster analysis. Comm. Stat. 3: 1–27.

Camarero, J. J., Gutiérrez, E. & Fortin, M.-J. 2000. Boundary detection in altitudinal treeline ecotones in the Spanish Central Pyrenees. Arct. Antarct. Alp. Res. 32: 117–126.

Chapin, T. G., Harrison, D. J. & Katnik, D. D. 1998. Influence of landscape pattern on habitat use by American marten in an industrial forest. Conserv. Biol. 12: 1327–37.

Cody, M. L. 1985. Habitat selection in birds. Academic Press, San Diego, CA.

Cohen, J. 1988. Statistical power for the behavioral sciences, 2nd edition. Academic Press, New York.

Collins, S. L., James, F. C. & Risser, P. G. 1982. Habitat relationships of wood warblers (Parulidae) in northern central Minnesota. Oikos 39: 50–58.

Csillag, F., Boots, B., Fortin, M.-J., Lowell, K. & Potvin, F. 2001. Multiscale characterization of boundaries and landscape ecological patterns. Geomatica, in press.

Fortin, M.-J. 1994. Edge detection algorithms for two dimensional ecological data. Ecology 75: 956–65.

Fortin, M.-J. 1997. Effects of data types on vegetation boundary delineation. Can. J. For. Res. 27: 1851–1858.

Fortin, M.-J. 1999. Effects of quadrat size and data measurement on the detection of boundaries. J. Veg. Sci. 10: 43–50.

Fortin, M.-J. & Drapeau, P. 1995. Delineation of ecological boundaries: Comparison of approaches and significance tests. Oikos 72: 323–332.

Fortin, M.-J., Drapeau, P. & Jacquez, G. M. 1996. Quantification of the spatial co-occurrences of ecological boundaries. Oikos 77: 51–60.

Freemark, K. E., & Merriam, H. G. 1986. Importance of area and habitat heterogeneity to bird assemblages in temperate forest fragments. Biol. Conserv. 36: 115–141.

Gates, F. C. 1942. The bogs of northern Lower Michigan. Ecol. Mon. 12: 213–254.

Gordon, A. D. 1999. Classification. 2nd edition. Monographs on Statistics and Applied Probability 82. Chapman & Hall/CRC Press, London.

Hall, K. R. 1996. Point counts of Michigan forest birds: Improving sampling efficiency for species of concern. Unpublished Master's Thesis, University of Michigan, Ann Arbor, MI.

120

Hansen, A. J. & diCastri, F. (eds). 1992. Landscape boundaries: Consequences for biotic diversity and ecological flows. Springer-Verlag, New York.

Hobbs, R. J. & Mooney, H. A. (eds). 1990. Remote sensing of biosphere functioning. Springer-Verlag, New York.

Jacquez, G. M. 1995. The map comparison problem: Tests for the overlap of geographic boundaries. Stat. Med. 14: 2343–2361.

Jacquez, G. M. & Maruca, S. 1998. Geographic boundary detection. In: Poiker, T. K. & Chrisman, N. (eds), Proceedings of the 8th internationational symposium on spatial data handling. International Geographical Union, Geographic Information Science Study Group, Burnaby, B.C., Canada.

Jacquez, G. M., Maruca, S. & Fortin, M.-J. 2000. From fields to objects: a review of geographic boundary analysis. J. Geogr. Sys. 2: 221–241.

James, F. C. & Shugart, H. H., Jr. 1970. A quantitative method of habitat description. Audubon Field Notes 24: 727–736.

James, F. C. & Wamer, N. O. 1982. Relationships between temperate forest bird communities and vegetation structure. Ecology 63: 159–171.

Johnston, C. A., Pastor, J., & Pinay, G. 1992. Quantitative methods for studying landscape boundaries. Pp. 107–125. In: Hansen, A. J. & diCastri, F. (eds), Landscape boundaries: Consequences for biotic diversity and ecological flows. Springer-Verlag, New York.

Küchler, A. W. 1967. Vegetation mapping. The Ronald Press Company, New York.

Legendre, L. & Legendre, P. 1983. Numerical Ecology. Developments in environmental modeling 3. Elsevier, Amsterdam.

Legendre, P. 1987. Constrained clustering. Pp. 289–307. In: Legendre, P. & Legendre, L. (eds), Developments in numerical ecology. NATO ASI Series, Vol. G14, Springer-Verlag, Berlin.

Legendre, P., & Fortin, M.-J. 1989. Spatial pattern and ecological analysis. Vegetatio 80: 107–138.

McGarigal, K. & McComb, W. C. 1995. Relationships between landscape structure and breeding birds in the Oregon coast range. Ecol. Monogr. 65: 235–60.

Milligan, G. W. & Cooper, M. C. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50: 159–179.

Millington, A. C. & Alexander, R. W. 2000. Vegetation mapping in the last three decades of the twentieth century. Pp. 321–331. In: Alexander, R. & Millington, A. C. (eds), Vegetation mapping: From patch to planet. John Wiley & Sons, New York.

Musick, H. B. & Grover, H. D. 1991. Image textural measures as indices of landscape pattern. Pp. 77–103. In: Turner, M. G. & Gardner, R. H. (eds), Quantitative methods in landscape ecology. Springer-Verlag, New York.

Oden, N. L., Sokal, R. R., Fortin, M.-J. & Goebl, H. 1993. Categorical wombling: detecting regions of significant change in spatially located categorical variables. Geogr. Anal. 25: 315–336.

Parody, J. M., Cuthbert, F. J. & Decker, E. H. In Press. The effect of 50 years of landscape change on species richness and community composition. Glob. Ecol. Biogeogr.

Parody, J. M. 1996. Avian assemblages in northern Michigan: A long-term perspective. Unpublished Master's Thesis, University of Michigan, Ann Arbor, MI.

Quattrochi, D. A. & Pelletier, R. E. 1991. Remote sensing for analysis of landscapes: an introduction. Pp. 51–76. In: Turner, M. G. & Gardner, R. H. (eds), Quantitative methods in landscape ecology. Springer-Verlag, New York.

Ralph, C. J. & Scott, J. M. (eds). 1981. Estimating the number of terrestrial birds. Studies in Avian Biol. 6.

Ralph, C. J., Geupel, G. R., Pyle, P., Martin, T. E. & DeSante, D. F. 1993. Handbook of field methods for monitoring landbirds. General Technical Report PSW-GTR-144. United States Department of Agriculture, United States Forest Service, Pacific Southwest Research Station, Albany, CA.

Scheaffer, R. L., Mendenhall, W. & Ott, L. 1990. Elementary survey sampling, 4th Edition. PWS-Kent Publishing Company, Boston.

Smith, K. G. 1977. Distribution of summer birds along a forest moisture gradient in an Ozark watershed. Ecology 58: 810–819.

Stohlgren, T. J., Owen, A. J. & Lee, M. Monitoring shifts in plant diversity in response to climate change: a method for landscapes. Biod. Cons. 9: 65–86.

Turner, S. J., O'Neill, R. V., Conley, W., Conley, M. R. & Humphries, H. C. 1991. Pattern and scale: statistics for landscape ecology. Pp. 17–49. In: Turner, M. G. & Gardner, R. H. (eds), Quantitative methods in landscape ecology. Springer-Verlag, New York.

Turner, M. G. & Gardner, R. H. (eds). 1991. Quantitative methods in landscape ecology. Springer-Verlag, New York.

Verner, J. 1985. Assessment of counting techniques. Pp. 247–302. In: Johnston, R. F. (ed.), Current ornithology, Vol. 2. Plenum Press, New York.

Villard, M.-A., Schmidt, E. V. & Maurer, B. A. 1998. Contribution of spatial modeling to avian conservation. Pp. 49–64. In: Marzluff, J. M. & Sallabanks, R. (eds), Avian conservation: research and management. Island Press, Washington, D.C.

White, P. S. & Pickett, S. T. A. 1985. Natural disturbance and patch dynamics: An introduction. Pp. 3–13. In: Pickett, S. T. A & White, P. S. (eds), The ecology of natural disturbance and patch dynamics. Academic Press, New York.

Wiens, J. A., Crawford, C. S. & Gosz, J. R. 1985. Boundary dynamics: a conceptual framework for studying landscape ecosystems. Oikos 45: 421–427.

Wiens, J. A. 1995. Landscape mosaics and ecological theory. Pp. 1–26. In: Hansson, L., Fahrig, L. & Merriam, G. (eds), Mosaic landscapes and ecological processes. Chapman & Hall, London.

Womble, W. H. 1951. Differential systematics. Science 114: 315–322.

Zogg, G. P. & B. V. Barnes. 1995. Ecological classification and analysis of wetland ecosystems, northern Lower Michigan, U.S.A. Can. J. For. Res. 25: 1865–1875.