

ON DETERMINING THE STATISTICAL SIGNIFICANCE OF DISCONTINUITIES WITHIN ORDERED ECOLOGICAL DATA¹

JOE M. CORNELIUS² AND JAMES F. REYNOLDS^{3,4}

Systems Ecology Research Group, San Diego State University, San Diego, California 92182-0401 USA

Abstract. Current ecological theory hypothesizes that boundaries between adjacent ecosystem units are important in determining ecosystem structure and function across heterogeneous landscapes, and that such boundaries are potentially important sites for early detection of global climate change effects. Hence, there is an increasing research effort to elucidate the structure and function of ecological boundaries. Yet traditional data analysis methods focus primarily on homogeneous units rather than on the boundaries between them; thus, new methods are being developed for detecting, characterizing and classifying boundaries, e.g., split moving-window boundary analysis (SMW). SMW is a simple yet sensitive method for locating discontinuities that may exist within multivariate, serial data (ordered in one dimension) at various scales relative to the length of the data series. However, SMW is subjective and relative, and therefore locates apparent discontinuities even within random, serial data. In this paper we present two nonparametric methods for determining the statistical significance of discontinuities detected by SMW. First, we describe a Monte Carlo method for determining the statistical significance of scale-dependent discontinuities (i.e., discontinuities that are significant relative to only one scale). Second, we propose a nonparametric, scale-independent method (it also is dependent upon scale size, but to a much lesser degree than the Monte Carlo method) that is more appropriate for locating statistically significant discontinuities that separate different, relatively homogeneous groups of varying size along a series. We examine the robustness of these two methods using computer-generated data having varying intensities of imposed discontinuities, and illustrate their application to locating boundaries between vegetation samples collected at systematic intervals across a desert landscape in southern New Mexico, USA.

Key words: *boundary detection; concentration contours; dissimilarity profile; gradient; Monte Carlo simulation; non-metric multidimensional scaling; nonparametric; ordination; permutation procedures; spatial pattern; statistical methods; transect.*

INTRODUCTION

The existence of boundaries, ecotones, or transition zones between relatively homogeneous ecosystem units has long been recognized by ecologists (e.g., Livingston 1903, Clements 1905), but interpretations of their ecological significance have been quite varied (cf. Leopold 1932, Whittaker 1956, van Leeuwen 1966, Daubenmire 1968, van der Maarel 1976, Margalef 1979, Allen and Starr 1982). Current landscape ecology theory hypothesizes that boundaries function in a fashion analogous to semi-permeable membranes, thereby controlling the rates of movement of abiotic and biotic components between units across heterogeneous landscapes (Wiens et al. 1985, Forman and Godron 1986). In addition, ecotones are hypothesized to be more dynamic relative to the larger homogeneous units they separate, and as such, may be important sites for the

early detection of global environmental change over a broad range of temporal and spatial scales (Hansen et al. 1988; however, see van der Maarel 1990). Given the potential significance of boundaries in current ecological theory and the paucity of information on them, more research is needed on elucidating the structure and function of ecosystem boundaries.

Most traditional methods for analyzing ecological data have concentrated on comparisons between homogeneous units, rather than on the boundaries between them. Thus, there is a need for further development of numerical methods for detecting, characterizing, and classifying boundaries (Wiens et al. 1985, Holland 1988, van der Maarel 1990). Quantitative methods for locating discontinuities in ordered data have been used by European phytogeographers and phytosociologists since the 1930s (van der Maarel 1976 and included references), and more recently by North American ecologists (e.g., Whittaker 1956, Ludwig and Cornelius 1987, and included references). The basic approach has been to calculate indices of compositional dissimilarity or similarity between successive pairs of ordered samples (e.g., vegetation stands ordered along field transects, or through ordination). Dissimilarity (or similarity) profile graphs are then con-

¹ Manuscript received 11 September 1990; revised 17 December 1990; accepted 28 December 1990.

² Present address: 914 East 42nd, San Angelo, Texas 76903 USA.

³ Address correspondence and reprint requests to this author.

⁴ Present address: Department of Botany, Duke University, Durham, North Carolina 27706 USA.

structed by plotting index scores vs. the midpoint locations between successive samples (also known as differential profiles, van der Maarel 1976, or distance profiles, Orlóci and Orlóci 1990). Sharp discontinuities in index values (i.e., peaks in a dissimilarity profile graph, troughs in a similarity profile graph) indicate the locations of relative discontinuities within the data series. A closely related method is "split moving-window boundary analysis," or SMW (Webster 1973, 1978, Ludwig and Cornelius 1987), which detects discontinuities within multivariate data ordered in one dimension through comparisons of dissimilarity metrics computed between the two halves of all sequential groups of samples (i.e., windows) of specific sizes; varying the window size (i.e., the number of ordered samples within each window) changes the scale at which boundaries are detected. The variable scale capability of SMW substantially improves the detection of discontinuities within noisy data.

Boundary detection methods are only useful for locating "relative" discontinuities (Whittaker 1956, Webster 1973). Consequently, SMW will "detect" discontinuities even within random data at both large and small scales relative to the length of the data series; hence, additional objective methods are necessary for determining statistical significance. In this paper, we present two objective methods for assessing the potential statistical significance of discontinuities identified by SMW. First, we describe an empirical, Monte Carlo method for determining the statistical significance of *scale-dependent* discontinuities (i.e., discontinuities that are significant relative to only a certain window size or scale). Second, we propose a nonparametric, *scale-independent* method (it also is dependent upon scale, but to a much lesser degree than the Monte Carlo method) that is more appropriate for locating statistically significant discontinuities that separate different, relatively homogeneous groups of widely varying size along a series. This method consists of hierarchical partitioning of a series using SMW results pooled from several scales, the calculation of test statistics from multi-response permutation procedures, MRPP (Mielke 1986, 1991), and examination of overlap in concentration contours between the groups defined at each level of partitioning within ordination diagrams from non-metric multidimensional scaling, NMDS (Davison 1983). We examine the robustness of these two methods using computer-generated data having varying intensities of imposed discontinuities, and illustrate their application to locating boundaries between vegetation samples collected at systematic intervals across a desert

ANALYTICAL METHODS

Boundary detection methods

A variety of boundary (or edge) detection methods (BDMs) have been used in the natural sciences (particularly geology, soil science, and ecology) and in the fields of robotics vision and biomedical digital imaging

landscape in southern New Mexico, USA.

(see Table 1). Some BDMs based on binary data are objective, utilizing statistics to locate boundaries (e.g., chi-square contingency analysis or Monte Carlo methods), while other approaches make use of parametric statistical comparisons between positions on both sides of all potential discontinuities (subject to the limiting assumptions of the test statistic used). However, most quantitative BDMs are essentially subjective; i.e., the researcher must decide if a discontinuity is "significant."

BDMs developed by soil scientists and geologists (i.e., SMW, global zonation, and the maximum level-variance method, MLV; Table 1) appear to have the greatest potential for general ecological application. However, global zonation and MLV are complex, multivariate statistical procedures based upon assumptions of independence and multivariate normality, and as such may not be appropriate when applied to ecological data that severely violate these assumptions (see also Turner et al. 1991). Webster (1978) found that SMW generally performs as well or better than MLV, and suggested that the simplicity and lower computational burden of SMW makes it more generally applicable. SMW is also the only BDM that directly incorporates variable scale, which is particularly useful with noisy data and for locating boundaries at different scales relative to the sample scale.

Split moving-window boundary analysis

Split moving-window boundary analysis (SMW) was developed by soil scientists for optimally locating boundaries between relatively homogeneous soil units along transects (Webster 1973, 1978). Our particular application of SMW consists of (refer to Appendix 1 for a mathematical description): (1) placing a window of a certain even-numbered size at the beginning of a data series, (2) splitting the window into two equal halves, (3) calculating averages for each variable within each half, (4) calculating some type of dissimilarity metric between the two halves of the window, (5) shifting the data window one position along the series, and (6) repeating steps 2–5 until the end of the data series is reached. Dissimilarity profile graphs are then constructed by plotting the dissimilarity values vs. the location of the window midpoint. Sharp peaks in dissimilarity suggest the locations of discontinuities. Ecological applications of SMW can be found in Ludwig and Cornelius (1987), Wierenga et al. (1987), Johnston et al. (1991), and Turner et al. (1991). See Brunt and Conley (1990) for a detailed study examining the behavior of SMW using squared euclidean distance on simulated boundaries of varying complexity.

Any one of a large number of dissimilarity measures could be used in SMW, depending upon whether the researcher wants to emphasize differences based on absolute abundances (e.g., euclidean distance), relative abundances (e.g., chord distance), or binary composi-

TABLE 1. Overview of quantitative methods for locating discontinuities in ordered data.

Method	Description	References
1. Compare indices of compositional similarity or dissimilarity between successive pairs of ordered samples	Plot index score vs. midpoint between adjacent samples, subjectively determine locations of discontinuities	van der Maarel 1976 and included references, Whittaker 1956, 1960, Beals 1969, Bratton 1975
2. Compare compositional similarity or dissimilarity of each sequential sample relative to an end point sample	Plot index score vs. sample position, subjectively determine positions where sharp transitions occur	van der Maarel 1976, Hobbs 1986
3. Compare species turnover rates along gradients	Plot ordination scores vs. sample location, subjectively determine positions where sharp transitions occur	Whittaker et al. 1979a, b, Shmida and Whittaker 1981, Hobbs 1986
4. Compare interspecific association between adjacent blocks of samples	Probabilistic determination of compositional discontinuities based on chi square	Lange and Sparrow 1985
5. Statistical analysis of overlap in species distributions	Probabilistic comparison of frequency of overlap in species distribution boundaries vs. expected value	Pielou and Routledge 1976, Pielou 1977, Gardiner and Haedrich 1978, Harper 1978, Underwood 1978, Dale 1986
6. Binary Monte Carlo estimation of expected number of species	Probabilistic comparison of number of species held in common between ordered samples relative to an expected number if all species were distributed at random	Raup and Crick 1979, McCoy et al. 1986
7. "Constrained" or "conditional" classification techniques	Maintain the order of samples in the formation of clusters	Gordon 1973, Lefkovitch 1980, Legendre et al. 1985, Legendre and Fortin 1989
8. Split moving-window boundary analysis (SMW)	Compare distances computed between the two halves of all sequential windows of specific sizes	Webster 1973, 1978, Ludwig and Cornelius 1987, Nwadiolo and Hole 1988
9. Global zonation procedure	Hierarchically divides series into segments that maximize the ratio of within-segment sums of squares to between-segment sums of squares	Gill 1970
10. Maximum level-variance method (MLV)	Searches for optimal partitioning of ordered data into a pre-specified number of segments	Hawkins and Merriam 1974
11. Orłóci edge detection method	Remove first-order serial effects; ordinate residuals with canonical contingency analysis; derive distance (edge location), deviations (edge intensity), and angles (edge sharpness) profiles for edge detection	Orłóci and Orłóci 1990
12. Other edge detection methods	From fields of robotics vision and biomedical digital imaging; primarily through differentiation or gradient measurements of optical intensity or density in one or two dimensions	Marr and Hildreth 1980 and included references, Smith et al. 1988

tion (e.g., Ochiai distance) (see Ludwig and Reynolds 1988). For ease of comparison, we use standard euclidean distance in all statistical methods presented below (i.e., SMW, MRPP, and NMDS).

The scale at which discontinuities are identified by SMW, and to some extent the type of discontinuity identified (e.g., sharp vs. gradual), is dependent upon the scale of the sampling and the window size used. Large window sizes are necessary to minimize the effects of noise on the results (i.e., minimize detecting a discontinuity that does not exist), and to maximize the ability to detect weak discontinuities (e.g., a trend dis-

continuity with a broad boundary, which may be undifferentiated from background noise at small scales). However, window sizes that are too large may not be able to detect discontinuities that exist at smaller scales (e.g., large windows that contain more than one discontinuity). Thus, results are sensitive to different window sizes, and selection of discontinuities usually requires the selection of an appropriate window size for the scale of interest.

Depending on objectives, the scale dependency of SMW may be either useful or confusing. Variation in the location and intensity of boundaries at different

scales may provide valuable information for research in landscape ecology, and for detecting subtle changes in ecosystem boundaries in response to environmental change. In such cases, Monte Carlo methods are most appropriate for determining the statistical significance of discontinuities. However, for other types of analyses (e.g., classification of vegetation zones along a transect), scale-related variability in the location and intensity of discontinuities may complicate problem solution. In these cases, the scale-dependency of SMW can be reduced by pooling information from several different scales, since the locations of peaks in a pooled dissimilarity profile are relatively scale independent compared to the dissimilarity profiles at each window size.

Determining the statistical significance of scale-dependent discontinuities

We used a Monte Carlo method to estimate the mean expected SMW dissimilarity and standard deviation between window halves for a given data series and window width, which we then used to determine the statistical significance of scale-dependent discontinuities under a null hypothesis that no discontinuities exist. Our method is related to that of Raup and Crick (1979) for boundary location with binary, species distribution data. Our Monte Carlo method consists of randomizing the position of each data vector along the series (which maintains the correlation structure of the data) and repeating the calculation of SMW dissimilarities as outlined previously. These randomized SMW calculations are repeated numerous times (1000 in our analyses), and then a mean dissimilarity and standard deviation are calculated for each window mid-point position (see Appendix 2). We then calculate an overall expected mean dissimilarity and standard deviation for a given window size by averaging the independent estimates from each position (see Appendix 3). We have examined the distributions of dissimilarity values from SMW analyses on several randomly generated data sets and several window sizes using euclidean distance, and almost all were normally distributed (Kolmogorov-Smirnov test statistic, $P < .05$ determining significance); hence, the mean and standard deviation are appropriate estimators of location and concentration.

We overlaid lines onto the dissimilarity profile graphs at values of 1, 2, and 3 standard deviations above the overall mean expected dissimilarity. We assumed that dissimilarity peaks extending above 2 standard deviations were significant for that particular data series and scale, which is a conservative enough criterion for determining statistical significance in most cases. Statistical confidence limits can also be defined around the expected mean dissimilarity using either standard parametric statistical methods (e.g., Steel and Torrie 1980), or an empirical approach similar to Orłóci and Beshir (1976). The Monte Carlo method is highly dependent upon scale size, and is relative only to the data series being analyzed, but can easily be

applied to SMW analyses using any type of dissimilarity measure.

Determining the statistical significance of scale-independent discontinuities

Hierarchical partitioning of the data series. — We first partitioned the series into groups using a pooled SMW dissimilarity profile (created by averaging together SMW dissimilarities from several different window sizes, for each window midpoint location). Because dissimilarities from small window sizes are usually greater than those from large window sizes, especially in very noisy data, simple averaging of dissimilarity values will often give more weight to the smaller window sizes. In order to give approximate equal weighting to dissimilarities from large and small scales, we transformed dissimilarity values for each window width into standardized variables, or Z scores, relative to the Monte Carlo estimate of the expected mean dissimilarity and standard deviation (Appendix 3), prior to averaging (Appendix 4). The resulting pooled dissimilarity profile is relatively scale-independent in comparison to the ones for each individual window size. We used this pooled dissimilarity profile to partition the data series into groups hierarchically. The location of the highest peak was used to partition the series into two groups (level 1). The second highest peak was used to partition one of the groups resulting from the first partition into two new groups (level 2). We continued partitioning each data series until there were no longer any visually discernible dissimilarity peaks.

Multi-response permutation procedures. — Multi-response permutation procedures (MRPP) represent a relatively new and powerful, distribution-free, multivariate statistical method for making comparisons between defined groups (Mielke 1986, 1991). MRPP is based on the within-group average of pairwise distance measures between object responses in a euclidean data space. While MRPP is functionally similar to MANOVA or Mahalanobis distance, MRPP may be more appropriate for most ecological data because it is permutation based; hence there are no distributional requirements of the data. In addition, MRPP is based on euclidean geometry, making it easy to relate the analysis to the perceived euclidean data space (Mielke 1991). See Biondini et al. (1985) and Zimmerman et al. (1985) for other ecological applications of MRPP.

MRPP provides a valuable statistical tool for determining the significance of potential discontinuities identified by SMW, because both methods can use the same dissimilarity measure, e.g., euclidean distance, chord distance, etc. We calculated MRPP test statistics and significance levels between the new groups formed at each level of hierarchical partitioning of the data series. We computed MRPP test statistics using the FORTRAN computer program MRPP (Mielke 1991) with the following options in effect: $v = 1$, which makes the analysis space of MRPP ordinary euclidean; $g = 2$,

for statistical comparisons between two defined groups; and $C(i) = n(i)/\Sigma n(i)$ for the within-group weighting factor, where $i = 1$ to g , and $n(i)$ is the number of objects within group i . (This weighting factor makes the MRPP test statistic (δ) the weighted average of within-group pairwise distance measure means.) We plotted MRPP significance levels vs. hierarchical partition level and examined the resulting graph for break-points, which we used along with significance levels as criteria for deciding the potential number of statistically significant discontinuities.

Nonmetric multidimensional scaling.—Nonmetric multidimensional scaling, NMDS (Davison 1983), is a nonlinear ordination method that gives an optimal, reduced-dimension, graphical representation of the relationships among observations in a data set. NMDS assumes no underlying structure to data other than that pairwise dissimilarities in the reduced-dimension space are monotonically related to dissimilarities in the original n -dimensional data space. Ecological applications of NMDS have shown that it causes the least distortion of original nonlinear data structure when compared to other common ordination methods (Kenkel and Orłóci 1986 and included references). Wartenberg et al. (1987) point out that NMDS may fail to order points correctly over very large gradients. However, this failure is likely due to failure of the particular dissimilarity measure to quantify differences between the ends of the gradient, and not due to failure of the NMDS algorithm. Bradfield and Kenkel (1987) adjusted pairwise dissimilarity values between points along large simulated gradients using a “flexible shortest path” algorithm, which greatly improved NMDS reconstruction of long gradients.

NMDS ordination of pairwise dissimilarities between sample units of a data series, with concentration contours (see following paragraph) graphed around the centroids of defined groups, provides a valuable graphical tool for determining the significance of potential discontinuities identified by SMW. Like SMW and MRPP, NMDS can easily use any dissimilarity measure. For each data set analyzed we calculated a matrix of pairwise euclidean distances between samples and conducted NMDS ordinations using BASIC programs modified from Ludwig and Reynolds (1988). We obtained 10 separate two-dimensional solutions for each data set, each with different random initial configurations. We selected the solution with the lowest minimum stress value as the optimal ordination solution. This lessened the possibility of obtaining NMDS solutions of local rather than global minima (Davison 1983).

Concentration contours.—We present results from NMDS ordinations by plotting the two-dimensional, rescaled axes vs. each other. We then compute and plot 1 standard deviation concentration contours (L. Orłóci, *personal communication*) around the centroid of each group by: (1) projecting all points within a group onto a single axis that passes through the group centroid

at set angles to the original axes; (2) calculating the mean and standard deviation for all points on the axis; and (3) mapping the location of the mean ± 1 standard deviation into the NMDS coordinate system. We then rotate the axis at 5° intervals for 180°, repeating steps 1–3 for each interval (see Appendix 5). These concentration contours serve as graphical representations of the effective size of a group of points, comparable to drawing “ellipses of equal concentration” around groups in an ordination diagram (Lagonegro and Feoli 1985), or confidence circles around group centroids in discriminant functions analysis (Pimentel 1983). The “ellipses of equal concentration” method assumes the cloud of points has a simple, bivariate normal distribution, and derives concentration ellipses from the variances around the group centroid in the two dimensions and the correlation between the two variates. Our method assumes only a normal distribution for the projection of all points onto a single axis. Thus the size of the standard deviation estimate can vary with the angle of the axis through the data cloud, which is more appropriate for defining concentration contours of data clouds that vary widely from a bivariate normal distribution. We preferred using 1 standard deviation contour widths to characterize group concentration, rather than constructing confidence sets on group means.

Data sets analyzed

Generated data.—We first examined the robustness of the methods described above using computer-generated data with varying intensities of discontinuity signals relative to the background noise. This objective was best met by generating data from simple uniform random distributions rather than from complex Gaussian distributions along gradients, which essentially have no boundaries. We examined five types of computer-generated data sets, all with 100 observations and 10 continuous variables. For each data set all 10 variables were generated from the same uniform random distribution. For data set No. 1, each variable for the entire series was generated from a range of 0–20. The other four types of generated data sets were similar to No. 1, except that step discontinuities were imposed at locations 34 and 67 of the series. For each of these generated data sets we varied the size intensity of the imposed step discontinuities relative to the background noise of the data. We indicate the intensity of an imposed discontinuity using a signal-to-noise ratio index from communication theory (Whalen 1971), in which signal size is divided by the root mean square noise (i.e., within-groups standard deviation). For example, we generated all variables from data set No. 2 from a range of 0–20 for locations 1–33, a range of 1–21 for locations 34–66, and a range of 2–22 for locations 67–100. A uniform random variable has variance defined as (maximum value – minimum value)²/12. Thus data set No. 2 had a discontinuity signal size of 1, a within-groups variance of 33.3333 [=20²/12], a root mean

TABLE 2. Characteristics of the computer-generated data sets analyzed. Each data set consisted of 10 variables and 100 observations, which were subdivided into three ranges: 1–33, 34–66, and 67–100. Each observation of all 10 variables within a range was generated from the same uniform random distribution.

Data set number	Range 1 1–33	Range 2 34–66	Range 3 67–100	Discontinuity step size	S/N ratio*
1	0–20	0–20	0–20	0	0.0000
2	0–20	1–21	2–22	1	0.1732
3	0–20	2–22	4–24	2	0.3464
4	0–20	4–24	8–28	4	0.6928
5	0–20	8–28	16–36	8	1.3856

* Signal-to-noise ratio. Calculation is described in *Analytical methods: Data sets analyzed: Generated data*.

square noise of 5.7735, and a signal-to-noise (S/N) ratio of 0.1732. Characteristics of the five generated data sets are summarized in Table 2.

Field data.—We also applied these methods to field data consisting of cover abundance estimates of vascular plant species taken from 89 sampling stations at 30-m intervals along a transect traversing a topographic gradient at the Jornada Long-Term Ecological Research (LTER) site (Wierenga et al. 1987, Orlóci and Orlóci 1990, Cornelius et al. 1991). The study site is located in the northern Chihuahuan Desert approximately 40 km north of Las Cruces, New Mexico, USA. The transect crosses a 1500-ha closed-basin watershed draining the slopes of an isolated granitic mountain (Mt. Summerford) and extends in a south-southwest direction from an ephemeral lake or playa located on the basin floor, across a fan piedmont and alluvial fan (bajada) located on the piedmont slope, and onto the base of the mountain (Cornelius et al. 1991). During mid-October of 1984 (the time of year when biomass accumulation from the summer growing season is at maximum), cover estimates of all vascular plant species were obtained by measuring the percentage of living crown cover intercepting 30-m lines positioned perpendicular to the main transect. For analysis we combined the cover estimates of individual plant species into guilds based upon a combination of photosynthetic pathway type (C_3 , C_4 , and CAM) and growth form (shrubs, subshrubs, succulents, and perennial and annual grasses and forbs; see Cornelius et al. 1991 for details). Nine vascular plant guilds occurred along the transect: C_3 shrubs, C_3 subshrubs, CAM succulents, C_4 perennial grasses, C_3 and C_4 perennial forbs, C_4 annual grasses, and C_3 and C_4 annual forbs.

RESULTS

Generated data

The scale of spatial separation between discontinuities relative to the length of the data series was known for the generated data sets; hence, we only include SMW results from window sizes of 10, 20, 30, and 40. In

addition, since results from data sets No. 1 (S/N ratio of 0.00) and No. 2 (S/N ratio of 0.1732) were similar, only the latter is presented. Likewise, results from data sets No. 4 (S/N ratio of 0.6928) and No. 5 (S/N ratio of 1.3856) were very similar, so only results for data set No. 4 are presented. All results are summarized as SMW euclidean distance profiles (Fig. 1), hierarchical partition levels (Fig. 2 and Table 3), and NMDS ordinations (Fig. 2). For illustrative purposes one of the 10 variables analyzed from each data set is presented (Fig. 1A, B, C, top panel). All 10 variables were used in the analyses (Fig. 1A, B, C, panels 2–6, Fig. 2).

In data set No. 2 (S/N ratio of 0.1732, Fig. 1A, panel 1), there were several sharp peaks in the SMW dissimilarity profiles (panels 2–5), but none occurred at the locations of the imposed step discontinuities (i.e., locations 33 and 66). At window sizes 10, 20, and 30 there were several statistically significant scale-dependent discontinuities, indicated by peaks >2 standard deviations above the Monte Carlo estimate of mean expected dissimilarity. Based on the pooled SMW dissimilarity profile (Table 3 and Fig. 1A, panel 6), we hierarchically partitioned the data series into three groups. Significance levels from MRPP test statistics between the new groups formed at the first two levels of partitioning were not significant (Fig. 2A, panel 1), and there was a large degree of overlap in concentration contours between the groups defined at each level of partitioning in the NMDS ordination diagrams (Fig. 2A, panels 2–4). These results indicate that even though the series may contain several statistically significant scale-dependent discontinuities, there are likely no significant discontinuities separating different homogeneous groups. Thus, the methods failed to detect the imposed discontinuities at this low level of S/N ratio.

In data set No. 3 (S/N ratio of 0.3464, Fig. 1B, panel 1), SMW detected discontinuities corresponding to the locations of the imposed step discontinuities (i.e., locations 33 and 66) for window sizes 20, 30, and 40 (but not at window size 10) (Fig. 1B, panels 2–6). However, only the peak at location 66 was >2 standard deviations above the mean expected dissimilarity. We made three hierarchical partitions of the data series based on the pooled SMW dissimilarity profile (Table

TABLE 3. Locations within each data series where potential discontinuities were identified at each hierarchical partition level based on the pooled split moving-window (SMW) dissimilarity profile.

Data set	Hierarchical partition level						
	1	2	3	4	5	6	7
1	55	38	20	73
2	30	10	81	47
3	63	32	9
4	32	65	21
5	32	66	47	11
Jornada	7	59	72	33	81	11	49

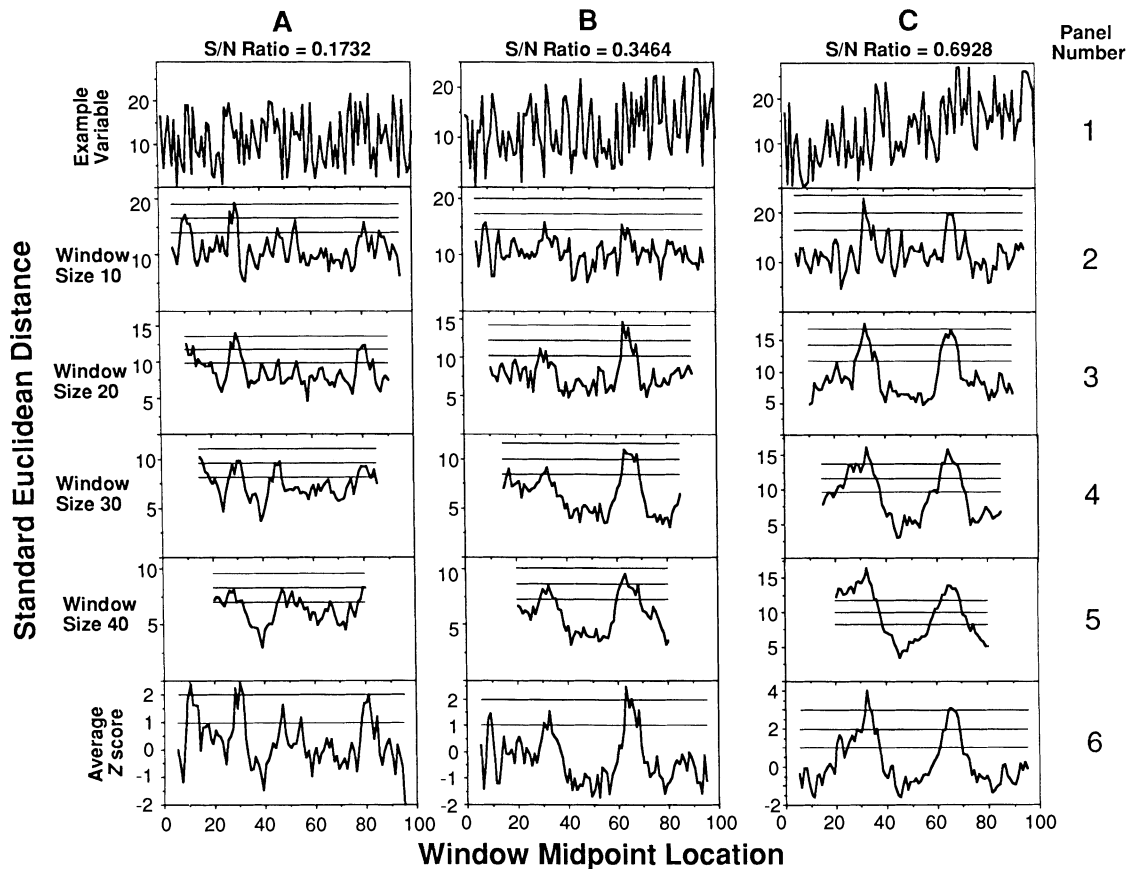


FIG. 1. Results from split moving-window (SMW) analyses of computer-generated data containing two step discontinuities with S/N ratios of (A) 0.1732, (B) 0.3464, and (C) 0.6928. Each of the three vertical sets of panels has graphs of window midpoint location vs. values for one of the 10 variables analyzed in each data set (panel 1), SMW dissimilarity profiles for window sizes 10, 20, 30, and 40 (panels 2–5), and the pooled average Z score profile (panel 6). Horizontal lines in the SMW dissimilarity profiles are at 1, 2, and 3 standard deviations above the overall expected mean euclidean distance.

3, Fig. 1B, panel 6). Only the significance level of MRPP test statistics between the new groups formed at the first partition level was significant (Fig. 2B, panel 1). The NMDS ordination diagram shows very little overlap between the groups defined at the first partition level, with progressively greater degrees of overlap at partition levels 2 and 3 (Fig. 2B, panels 2–4). Thus, our analysis reveals that there are likely one or two statistically significant discontinuities separating relatively homogeneous groups within the data series.

In data set No. 4 (S/N ratio of 0.6928, Fig. 1C, panel 1), SMW showed prominent peaks at locations 33 and 66 at all window sizes 10–40 (Fig. 1C, panels 2–5); and the SMW peaks were all >2 standard deviations above the mean expected dissimilarity. We made four hierarchical partitions of the data series based on the pooled dissimilarity profile (Table 3, Fig. 1C, panel 6). Significance levels from MRPP test statistics between the new groups formed at each partition level indicate that as many as three levels could be statistically significant (Fig. 2C, panel 1). The NMDS ordination diagram

shows no overlap of concentration contours between groups at partition levels 1 and 2, with substantial overlap occurring at partition level 3 (Fig. 2C, panels 2–4). Our analysis revealed that there were likely two or three statistically significant discontinuities separating relatively homogeneous groups within the data series.

Results from data set No. 5 (S/N ratio of 1.3896, not shown) revealed the presence of two statistically significant discontinuities within the data series.

Field data

The scale of field sampling restricted our SMW analyses to small window sizes (i.e., 4–12). Dissimilarity profiles from small window sizes (e.g., 4 and 6) were quite noisy in comparison to those from larger ones (e.g., 8, 10, and 12), but the dissimilarity peaks at transect positions 7 and 59 were prominent at even small window sizes (Fig. 3A, panels 2–6). Only at locations 7 (for window sizes 4–12) and 59 (for window sizes 10–12) were the discontinuities greater than 2 standard

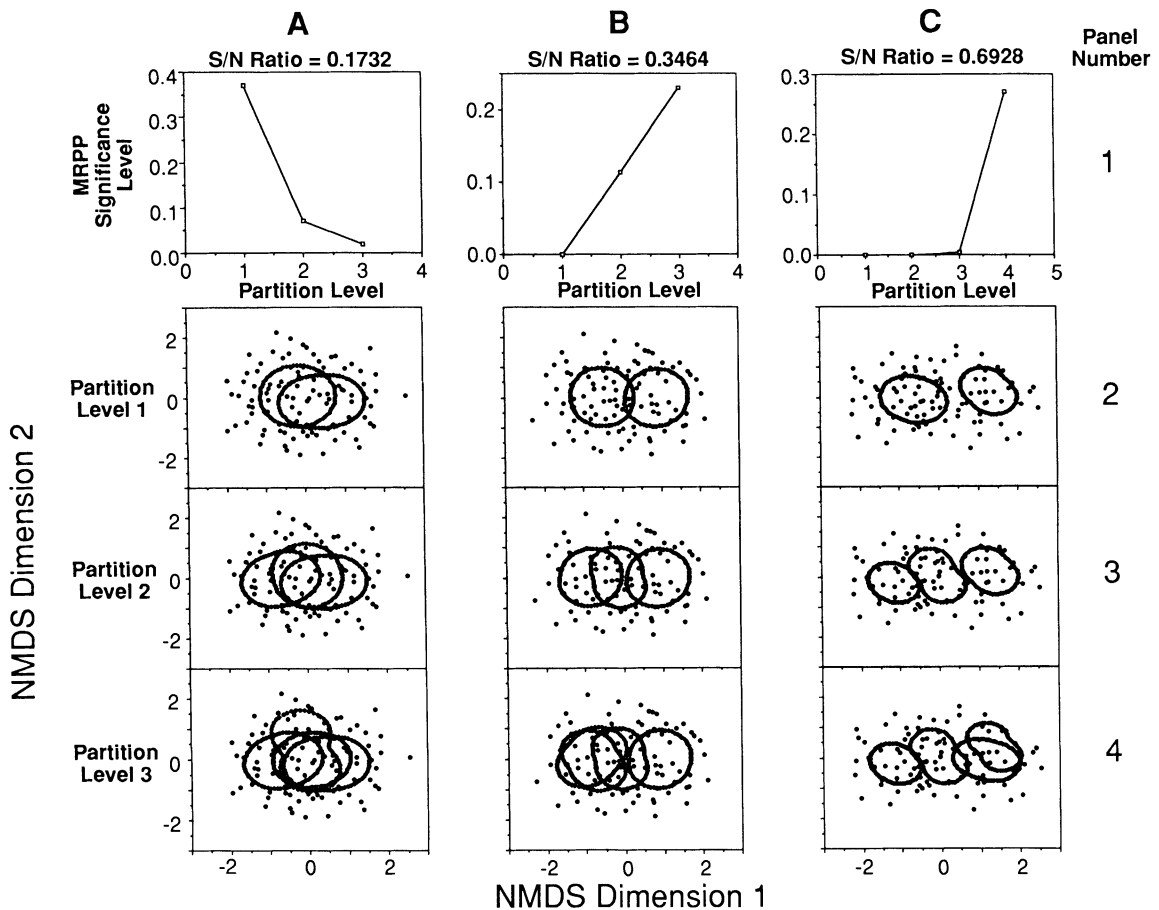


FIG. 2. Results from MRPP and NMDS analyses of computer-generated data containing two step discontinuities and S/N ratios of (A) 0.1732, (B) 0.3464, and (C) 0.6928. Each vertical panel has plots of hierarchical partition level vs. significance levels from MRPP test statistics for comparisons between the new groups formed at each level (panel 1, top); and NMDS ordination diagrams with one standard deviation concentration contours around the centroids of all groups formed at hierarchical partition levels 1–3 (panels 2–4).

deviations above the expected mean dissimilarity. We made seven hierarchical partitions of the data series based on the pooled SMW dissimilarity profile (Table 3, Fig. 3A, panel 7). MRPP significance levels indicate that up to six partitioning levels are likely significant (Fig. 3A, panel 1). NMDS ordination diagrams showed no overlap in concentration contours around the centroids of groups defined at partition levels 1–5 (Fig. 3B, panels 1–5).

Overall, our objective analysis suggests the presence of six statistically significant discontinuities separating seven homogeneous groups in the Jornada vegetation spatial series, with locations corresponding closely to vegetation zones identified subjectively in previous studies (Ludwig and Cornelius 1987, Wierenga et al. 1987, Orlóci and Orlóci 1990, Cornelius et al. 1991). Average cover of the nine vascular plant guild types varied tremendously between the seven major vegetation zones (Table 4), with C_3 shrubs dominating the playa fringe and bajada shrubland zones, C_4 perennial

grasses dominating the playa, lower and upper piedmont grassland zones, C_3 annual forbs dominating the upper mixed basin slopes zone, and C_3 annual forbs and C_4 perennial grasses co-dominating the lower mixed basin slopes zone. Previous studies have shown that these vascular plant guilds are apparently distributed along the transect relative to a complex environmental gradient of available water and nitrogen (Cornelius et al. 1991).

DISCUSSION

Our results indicate that locating statistically significant discontinuities within a data series is difficult, even with the aid of objective methods. This is especially true when a series contains very subtle discontinuities (e.g., the discontinuities in the generated data sets with S/N ratios of 0.1732 and 0.3464). Sharp, obvious boundaries are easily differentiated by the quantitative methods, but they are also easily determined subjectively (e.g., the imposed discontinuities

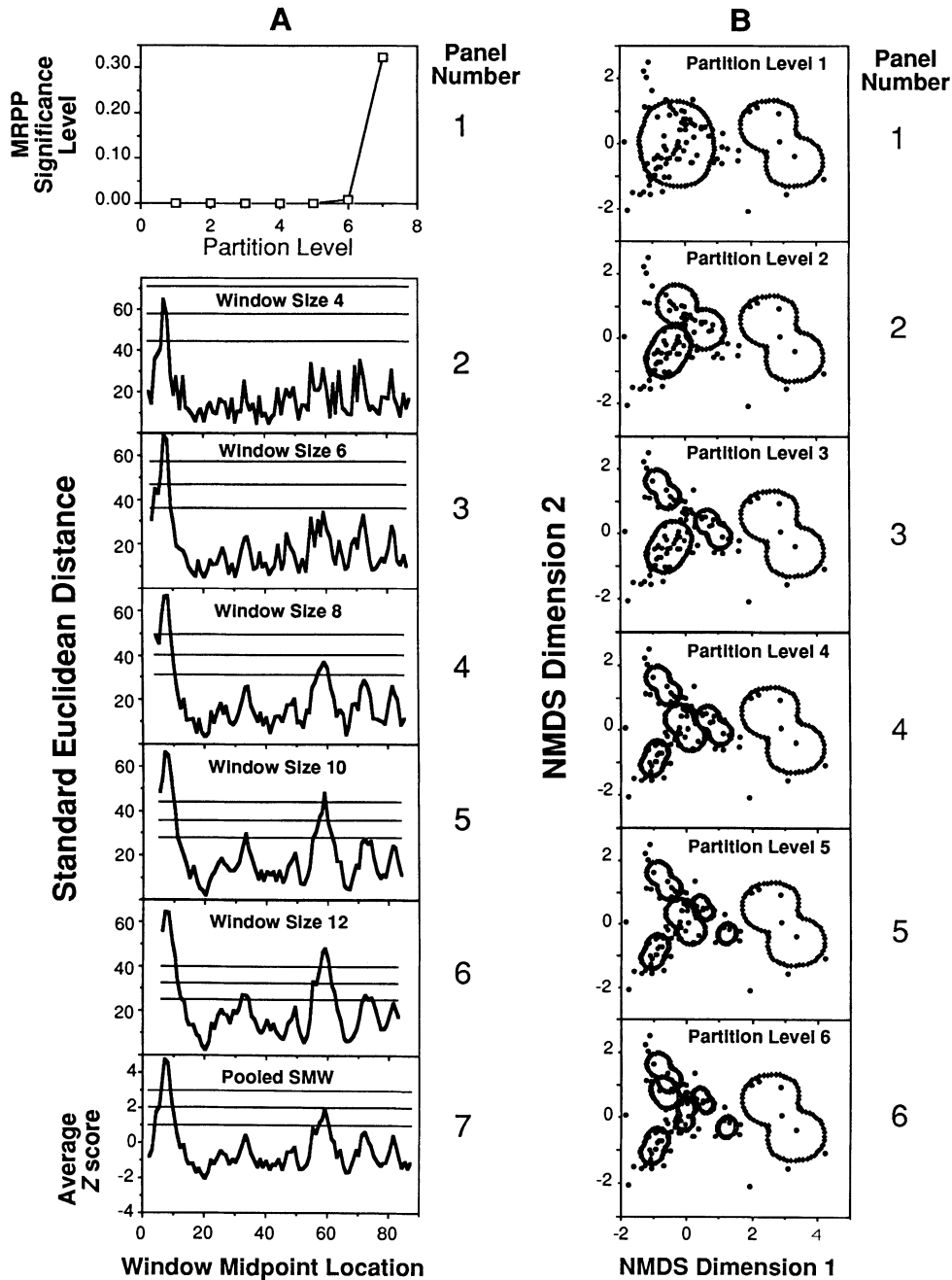


FIG. 3. Results from analyses of the LTER Jornada vegetation spatial series. (A) MRPP significance level from comparisons between the new groups formed at each hierarchical partition level (panel 1, top), SMW dissimilarity profiles for window sizes 4, 6, 8, 10, and 12 (panels 2–6), and for the average Z-score of window sizes 4–12 (panel 7). (B) NMDS ordination diagrams with 1 standard deviation concentration contours around the centroids of groups formed at hierarchical partition levels 1–6 (panels 1–6). Horizontal lines in the SMW dissimilarity profiles in part A are at 1, 2, and 3 standard deviations above the overall expected mean euclidean distance.

within the generated data sets with S/N ratios of 0.6928 and 1.3896, and the boundaries at transect positions 7 and 59 within the field data set). Objective, quantitative methods are necessary for determining the statistical significance of discontinuities, but in many applications may be most useful for defining lower and

upper limits to the potential number of significant discontinuities rather than defining their actual number.

We used a nonparametric, Monte Carlo method for determining the statistical significance of scale-dependent discontinuities. Their significance could also be determined by calculating some type of multivariate

TABLE 4. Average fall 1984 cover of vascular plant guilds within each of the seven significant vegetation zones identified by split moving-window boundary analysis (SMW). Names of vegetation zones are adapted from Wierenga et al. (1987) and Cornelius et al. (1990).

Vegetation zone	Transect stations	C ₃ shrubs	C ₃ sub-shrubs	CAM succulents	C ₄ perennial grasses	C ₃ perennial forbs	C ₄ perennial forbs	C ₄ annual grasses	C ₃ annual forbs	C ₄ annual forbs
Playa	1-7	0.0	0.0	0.0	62.6	39.1	0.0	0.0	1.6	3.6
Playa fringe	8-11	24.9	5.8	0.0	21.0	5.5	2.8	0.0	1.3	20.2
Lower mixed basin slopes	12-33	1.3	5.7	0.1	14.0	9.4	3.2	3.4	2.5	16.4
Upper mixed basin slopes	34-59	2.5	4.4	0.0	6.4	5.5	1.8	9.3	1.7	49.3
Bajada shrubland	60-72	25.5	3.2	0.5	6.0	3.2	0.7	0.0	0.0	3.6
Lower piedmont grassland	73-81	4.9	3.9	2.9	23.2	4.2	0.9	9.0	5.4	1.8
Upper piedmont grassland	82-89	4.1	0.6	3.6	42.1	1.3	3.8	5.0	3.2	9.1

test statistic between the two window halves, such as in the original implementation by soil scientists, e.g., Hotelling's T^2 (Webster 1973, 1978), Lawley-Hotelling trace statistic (Wierenga et al. 1987) or Wilk's lambda (Nwadialo and Hole 1988). However, this approach is not ideal because it requires stringent statistical assumptions, it makes comparisons between the two halves of a *window* and not between defined *groups* (i.e., it is highly scale dependent), and the number of observations must always be greater than the number of variables (not always possible with ecological data that have large numbers of species and other variables). Webster used principal components analysis (PCA) to reduce the dimensionality of the data space, and calculated dissimilarities between window halves using scores from the first few principal axes. However, PCA assumes linear relations among data and should only be used when this assumption is not severely violated (Orlóci 1979). Thus, for many ecological data Webster's approach may not be appropriate.

The Monte Carlo method successfully located all discontinuities in the generated data series with high S/N ratios (0.6928 and 1.3896), but failed to identify all boundaries in the low S/N ratio data series (0.1732 and 0.3464). This is because the definition of limits for determining the statistical significance of dissimilarity peaks is dependent upon the variation present within the *entire* series. In a highly variable series this could lead to an erroneous determination of nonsignificance for boundaries that may actually be significant (i.e., a type II statistical error). For example, for any given window size, analysis of the field data set with the Monte Carlo method identified only one or two boundaries as statistically significant, while the scale-independent method identified six boundaries as potentially significant.

The extreme scale dependency of SMW may cause problems when analyzing complex data series consisting of several boundaries separated by widely varying spatial distances (such as the case for the field data set we analyzed). We addressed this problem by combining information from several scale sizes into a single dissimilarity profile, which we then used to partition

the data series into groups hierarchically. Other methods could also be used to address the scale-dependency problem of SMW. For example, Webster (1973) used autocorrelation analysis of ordered scores on the first few principal components to estimate an average spatial distance between boundaries along a data series, then chose a window size of $\approx 2/3$ this average distance as optimal. Similarly, optimal window sizes could be determined using semivariogram analysis (e.g., Robertson 1987), quadrat variance spatial analysis methods (e.g., Gibson and Greig-Smith 1986, Ver Hoef and Glenn-Lewin 1989), and multivariate geostatistical methods (e.g., Legendre and Fortin 1989). However, determining an optimal SMW window size in this manner assumes that all boundaries are close to the same distance apart, and will only be appropriate when this assumption is not severely violated.

We used nonparametric methods for determining the statistical significance of scale-independent discontinuities defined from hierarchical partitioning of a pooled, SMW dissimilarity profile. One approach consisted of calculating test statistics from MRPP between the new groups defined at each level of partitioning. Traditionally, multivariate comparisons between a priori defined groups have been made using multivariate analysis of variance (MANOVA, for multi-group comparisons), Mahalanobis distance and its associated Hotelling's T^2 test statistic (for pairwise group comparisons), and discriminant functions analysis (e.g., Klecka 1980, Pimentel 1983). Proper use of these methods requires meeting assumptions of multivariate normality and equality of covariance matrices, assumptions often violated with ecological data. MRPP is permutation based, so there are no distributional requirements of the data, which makes it more appropriate for many ecological applications (Mielke 1986, 1991).

In addition to MRPP, we also examined NMDS ordination diagrams and the overlap in confidence contours between the centroids of new groups defined by each level of partitioning. Such graphical analyses of relationships among sample units in ecological data have traditionally been made using dimension-reduc-

tion ordination techniques based on eigenanalysis (Orlóci 1978, Legendre and Legendre 1983, Ludwig and Reynolds 1988). These ordination methods (e.g., principal components analysis, correspondence analysis or reciprocal averaging, and metric multidimensional scaling or principal coordinates analysis) assume linear relationships among variables, an assumption violated by many ecological data (Beals 1973, Orlóci 1979). Other ordination methods (e.g., Gaussian ordination and detrended correspondence analysis) have been developed to "adjust" for nonlinearity of data, but they also stringently assume an underlying structure to the data. NMDS is a nonlinear ordination method that gives an optimal, reduced-dimension, graphical representation of the relationships among observations in a data set, and that assumes no underlying structure to the data other than that pairwise dissimilarities in the reduced dimensional space are monotonically related to dissimilarities in the original n -dimensional data space. Thus, NMDS is more appropriate for ordination of most ecological data (see also Kenkel and Orlóci 1986).

One appealing advantage for using MRPP and NMDS to determine the significance of discontinuities identified by SMW is that all methods can use the same dissimilarity measure. The examples we presented all used standard euclidean distance, which makes comparisons between absolute abundances of variables, and permits easy geometric interpretation of results. However, if data vectors are normalized prior to analysis, then results will be based on chord distance, or the differences in relative abundances of the variables. NMDS and SMW can easily use any dissimilarity metric. The particular computer program we used for calculating MRPP test statistics (Mielke 1991) is based only on euclidean distance type measures, because their behavior is understood enough so that a full permutation estimate is unnecessary. There is also a full permutation version of MRPP that can use any dissimilarity measure, but it is extremely computationally intensive for large-sample sizes. Thus, the methods we have applied here can be modified for probabilistic partitioning of serial data using any type of qualitative or quantitative dissimilarity measure.

In summary, we proposed nonparametric methods for determining the statistical significance of both scale-dependent and scale-independent discontinuities that may exist within a data series, as identified by SMW. The scale-dependent method compares the peaks in SMW dissimilarities relative to a Monte Carlo estimate of the mean dissimilarity and standard deviation expected for a given series. The scale-independent method consists of hierarchical partitioning of a series using SMW results pooled from several scale sizes, and the calculation of MRPP test statistics and examination of overlap in concentration contours in NMDS ordination diagrams between the groups defined at each level of partitioning. These methods provide valuable, ob-

jective criteria for determining the potential number of statistically significant discontinuities within a data series. However, even the simple examples included here showed that the methods are not ideal, as they sometimes indicated the potential significance of one discontinuity more or less than was actually imposed in the generated data (type I and type II errors inherent to all statistical procedures). However, our methods do allow definition of lower and upper limits to the potential number of statistically significant discontinuities that may exist, which is valuable information for interpreting complex, multivariate data from a long spatial or temporal series.

ACKNOWLEDGMENTS

This research was partially supported by grants from the DOE, Carbon Dioxide Research Division (DE-FG603-86 ER60490) and the NSF, Long-Term Ecological Research Program (BSR-88-11160). We acknowledge László Orlóci for suggesting the calculation of confidence contours. We thank Kenneth Burnham, Paul Kemp, Norman Kenkel, László Orlóci, and Solange Silva for critical reviews of earlier versions of this manuscript.

LITERATURE CITED

- Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: perspectives for ecological complexity. University of Chicago Press, Chicago, Illinois, USA.
- Beals, E. W. 1969. Vegetational change along altitudinal gradients. *Science* **165**:981-985.
- . 1973. Ordination: mathematical elegance and ecological naivete. *Journal of Ecology* **61**:23-35.
- Biondini, M. E., C. D. Bonham, and E. F. Redente. 1985. Secondary successional patterns in a sagebrush (*Artemisia tridentata*) community as they relate to soil disturbance and soil biological activity. *Vegetatio* **60**:25-36.
- Bradfield, G. E., and N. C. Kenkel. 1987. Nonlinear ordination using flexible shortest path adjustment of ecological distances. *Ecology* **68**:750-753.
- Bratton, S. P. 1975. A comparison of beta diversity functions of the overstory and herbaceous understory of a deciduous forest. *Bulletin of the Torrey Botanical Club* **102**:55-60.
- Brunt, J. W., and W. Conley. 1990. Behavior of a multivariate algorithm for ecological edge detection. *Ecological Modelling* **49**:179-203.
- Clements, F. E. 1905. Research methods in ecology. Universal, Lincoln, Nebraska, USA.
- Cornelius, J. M., P. R. Kemp, J. A. Ludwig, and G. L. Cunningham. 1991. The distribution of vascular plant species and guilds in space and time along a desert gradient. *Journal of Vegetation Science* **2**:59-71.
- Dale, M. R. T. 1986. Overlap and spacing of species ranges on an environmental gradient. *Oikos* **47**:303-308.
- Daubenmire, R. 1968. Plant communities. Harper, New York, New York, USA.
- Davison, M. L. 1983. Multidimensional scaling. John Wiley & Sons, New York, New York, USA.
- Forman, R. T. T., and M. Gordon. 1986. Landscape ecology. John Wiley & Sons, New York, New York, USA.
- Gardiner, F. P., and R. L. Haedrich. 1978. Zonation in the deep benthic megafauna: application of a general test. *Oecologia (Berlin)* **31**:311-317.
- Gibson, D. J., and P. Greig-Smith. 1986. Community pattern analysis: a method for quantifying community mosaic structure. *Vegetatio* **66**:41-47.

- Gill, D. 1970. Application of a statistical zonation method to reservoir evaluation and digitized log analysis. *Bulletin of the American Association of Petroleum Geology* **54**:719–729.
- Gordon, A. D. 1973. Classification in the presence of constraints. *Biometrics* **29**:821–827.
- Hansen, A. J., F. di Castri, and R. J. Naiman. 1988. A new SCOPE project. Ecotones in changing environments: the theory and management of landscape boundaries. *Biology International, Special Issue* **17**:137–161.
- Harper, C. W. 1978. Groupings by locality in community ecology and paleoecology: tests of significance. *Lethaia* **11**:251–257.
- Hawkins, D. M., and D. F. Merriam. 1974. Zonation of multivariate sequences of digitized geologic data. *Journal of the International Association of Mathematical Geology* **6**:263–269.
- Hobbs, E. R. 1986. Characterizing the boundary between California annual grassland and coastal sage scrub with differential profiles. *Vegetatio* **65**:115–126.
- Holland, M. M., compiler. 1988. SCOPE/MAB technical consultations on landscape boundaries: report of a SCOPE/MAB workshop on ecotones 5–7 January 1987, Paris, France. *Biology International, Special Issue* **17**:47–106.
- Johnston, C. A., J. Pastor, and G. Pinay. 1991. Quantitative methods for studying landscape boundaries. *In* F. di Castri and A. Hansen, editors. *Landscape boundaries: consequences for biotic diversity and ecological flows*. Springer-Verlag, New York, New York, USA, *in press*.
- Kenkel, N. C., and L. Orłóci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. *Ecology* **67**:919–928.
- Klecka, W. R. 1980. *Discriminant analysis*. Sage University Paper series: quantitative applications in the Social Sciences, Series Number 07-019. Sage, Beverly Hills, California, USA.
- Lagonegro, M., and E. Feoli. 1985. The use of ellipses of equal concentration to analyse ordination vegetation patterns. *Studia Geobotanica* **5**:143–165.
- Lange, R. T., and A. D. Sparrow. 1985. Moving analysis of interspecific associations. *Australian Journal of Botany* **33**:639–644.
- Lefkovich, L. P. 1980. Conditional clustering. *Biometrics* **36**:43–58.
- Legendre, L., and P. Legendre. 1983. *Numerical ecology*. Elsevier Scientific, New York, New York, USA.
- Legendre, P., S. Dallot, and L. Legendre. 1985. Succession of species within a community: chronological clustering, with applications to marine and freshwater zooplankton. *American Naturalist* **125**:257–288.
- Legendre, P., and J. Fortin. 1989. Spatial pattern and ecological analysis. *Vegetatio* **80**:107–138.
- Leopold, A. 1932. *Game management*. Scribner's Sons, New York, New York, USA.
- Livingston, B. E. 1903. The distribution of the upland societies of Kent County, Michigan. *Botanical Gazette* **35**:36–55.
- Ludwig, J. A., and J. M. Cornelius. 1987. Locating discontinuities along ecological gradients. *Ecology* **68**:448–450.
- Ludwig, J. A., and J. F. Reynolds. 1988. *Statistical ecology*. John Wiley & Sons, New York, New York, USA.
- Margalef, R. 1979. The organization of space. *Oikos* **33**:152–159.
- Marr, D., and E. Hildreth. 1980. Theory of edge detection. *Proceedings of the Royal Society of London, Series B* **207**:187–217.
- McCoy, E. D., S. S. Bell, and K. Walters. 1986. Identifying biotic boundaries along environmental gradients. *Ecology* **67**:749–759.
- Mielke, P. W. 1986. Non-metric statistical analyses: some metric alternatives. *Journal of Statistical Planning and Inference* **13**:377–387.
- . 1991. The application of multivariate permutation methods based on distance functions in the earth sciences. *Earth-Science Reviews, in press*.
- Nwadialo, B. E., and F. D. Hole. 1988. A statistical procedure for partitioning soil transects. *Soil Science* **145**:58–62.
- Orłóci, L. 1978. *Multivariate analysis in vegetation research*. Second edition. Junk, The Hague, The Netherlands.
- . 1979. Non-linear data structures and their description. Pages 191–202 *in* L. Orłóci, C. R. Rao, and W. M. Stiteler, editors. *Multivariate methods in ecological work*. International Co-operative Publishing House, Fairland, Maryland, USA.
- Orłóci, L., and M. E. Beshir. 1976. A heuristic test for homogeneity in species populations. *Vegetatio* **31**:141–145.
- Orłóci, L., and M. Orłóci. 1990. Edge detection in vegetation: Jornada revisited. *Journal of Vegetation Science* **1**:311–324.
- Pielou, E. C. 1977. The latitudinal spans of seaweed species and their patterns of overlap. *Journal of Biogeography* **4**:299–311.
- Pielou, E. C., and R. D. Routledge. 1976. Salt marsh vegetation: latitudinal gradients in the zonation patterns. *Oecologia (Berlin)* **24**:311–321.
- Pimentel, R. A. 1983. *Morphometrics. The multivariate analysis of biological data*. Kendall/Hunt, Dubuque, Iowa, USA.
- Raup, D. M., and R. E. Crick. 1979. Measurement of faunal similarity in paleontology. *Journal of Paleontology* **53**:1213–1227.
- Robertson, G. P. 1987. Geostatistics in ecology: interpolating with known variance. *Ecology* **68**:744–748.
- Shmida, A., and R. H. Whittaker. 1981. Pattern and biological microsite effects in two shrub communities, southern California. *Ecology* **62**:234–251.
- Smith, T. G., Jr., W. B. Marks, G. D. Lange, W. H. Sheriff, Jr., and E. A. Neale. 1988. Edge detection in images using Marr-Hildreth filtering techniques. *Journal of Neuroscience Method* **26**:75–82.
- Steel, R. G. D., and J. H. Torrie. 1980. *Principles and procedures of statistics: a biometrical approach*. Second edition. McGraw-Hill, New York, New York, USA.
- Turner, S. J., R. V. O'Neill, W. Conley, M. R. Conley, and H. C. Humphries. 1991. Pattern and scale: statistics for landscape ecology. Pages 17–49 *in* M. G. Turner and R. H. Gardner, editors. *Quantitative methods in landscape ecology*. Springer-Verlag, New York, New York, USA.
- Underwood, A. J. 1978. The detection of non-random patterns of distribution of species along a gradient. *Oecologia (Berlin)* **36**:317–326.
- van der Maarel, E. 1976. On the establishment of plant community boundaries. *Berichte der Deutschen Botanischen Gesellschaft* **89**:415–443.
- . 1990. Ecotone and ecoclines are different. *Journal of Vegetation Science* **1**:135–138.
- van Leeuwen, C. G. 1966. A relation-theoretical approach to pattern and process in vegetation. *Wentia* **15**:25–41.
- Ver Hoef, J. M., and D. C. Glenn-Lewin. 1989. Multiscale ordination: a method for detecting pattern at several scales. *Vegetatio* **82**:59–67.
- Wartenberg, D., S. Ferson, and F. J. Rohlf. 1987. Putting things in order: a critique of detrended correspondence analysis. *American Naturalist* **129**:434–448.
- Webster, R. 1973. Automatic soil-boundary location from transect data. *Journal of the International Association of Mathematical Geology* **5**:27–37.
- . 1978. Optimally partitioning soil transects. *Journal of Soil Science* **29**:388–402.

Whalen, A. D. 1971. Detection of signals in noise. Academic Press, New York, New York, USA.
 Whittaker, R. H. 1956. Vegetation of the Great Smoky Mountains. *Ecological Monographs* 26:1-80.
 ———. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279-338.
 Whittaker, R. H., L. E. Gilbert, and J. H. Connell. 1979a. Analysis of two-phase pattern in a mesquite grassland, Texas. *Journal of Ecology* 67:935-952.
 Whittaker, R. H., W. A. Niering, and M. D. Crisp. 1979b. Structure, pattern and diversity of a mallee community in New South Wales. *Vegetatio* 39:65-76.

Wiens, J. A., C. Crawford, and J. R. Gosz. 1985. Boundary dynamics as a conceptual framework for studying landscape ecosystems. *Oikos* 45:421-427.
 Wierenga, P. J., J. M. H. Hendrickx, M. H. Nash, J. A. Ludwig, and L. A. Daugherty. 1987. Variation of soil and vegetation with distance along a transect in the Chihuahuan Desert. *Journal of Arid Environments* 13:53-63.
 Zimmerman, G. M., H. Goetz, and P. W. Mielke, Jr. 1985. Use of an improved statistical method for group comparisons to study effects of prairie fire. *Ecology* 66:606-611.

APPENDIX 1

CALCULATION OF SPLIT MOVING-WINDOW DISSIMILARITIES

Consider a data series of $j = 1, 2, 3, \dots, N$ ordered positions with $i = 1, 2, 3, \dots, V$ measurement variables at each position. X_{ij} is the abundance of the i th variable at the j th position along the series. Consider a window of width Q that brackets sequential positions along the series, as in the calculation of moving averages. A series of length N will have $N-Q$ sequential windows of width Q . If Q is a non-zero, even integer $< N$, then a window of width Q can be split into two equal halves, designated by W_A and W_B , with each half window having $Q/2$ sequential positions. The location of each window can be uniquely defined by the location of its window midpoint $k + 0.5$, where $k = Q/2, Q/2 + 1, Q/2 + 2, \dots, N-Q/2$. For each window midpoint location the average of each variable i in each window half is given by

$$\bar{W}_{Ak+0.5,i} = \frac{\sum_{j=k-Q/2}^k X_{ij}}{Q/2}$$

and

$$\bar{W}_{Bk+0.5,i} = \frac{\sum_{j=k+1}^{k+Q/2} X_{ij}}{Q/2}$$

A dissimilarity/similarity index ($DS_{k+0.5}$) can be calculated between each of the resulting $N-Q$ pairs of average vectors. For each window midpoint location the standard euclidean distance between average, half-window vectors is given by

$$DS_{k+0.5} = \left[\sum_{i=1}^V (\bar{W}_{Ak+0.5,i} - \bar{W}_{Bk+0.5,i})^2 \right]^{1/2}$$

APPENDIX 2

CALCULATION OF OVERALL EXPECTED MEAN DISSIMILARITY AND STANDARD DEVIATION

Using Monte Carlo simulation techniques an expected mean dissimilarity can be estimated for an SMW dissimilarity array calculated for a given window width Q . Each data vector is randomly repositioned along the data series for $l = 1, 2, 3, \dots, M$ times, and SMW dissimilarities (as in Appendix 1) are calculated for each of the re-ordered data sets, resulting in an array of dissimilarities, $DR_{k+0.5,l}$. The mean expected dissimilarity and standard deviation for each $k + 0.5$ window midpoint location along the series is given by

$$\overline{DS}_{k+0.5} = \frac{\sum_{l=1}^M DR_{k+0.5,l}}{M}$$

and

$$SD_{k+0.5} = \frac{\left[\sum_{l=1}^M (DR_{k+0.5,l} - \overline{DS}_{k+0.5})^2 \right]^{1/2}}{M - 1}$$

and the overall expected mean dissimilarity and average standard deviation for the series at window width Q is given by

$$\overline{DS}_* = \frac{\sum_{k=Q/2}^{N-Q/2} \overline{DS}_{k+0.5}}{N - Q}$$

and

$$\overline{SD} = \frac{\sum_{k=Q/2}^{N-Q/2} SD_{k+0.5}}{N - Q}$$

APPENDIX 3

Z SCORE TRANSFORMATION OF SMW DISSIMILARITIES

For each window midpoint location the SMW dissimilarity estimate from a window of a given width can be transformed to a standardized variable, or Z score ($DZ_{k+0.5}$), relative to the overall expected mean dissimilarity and standard deviation for that window width (from Appendix 2) by

$$DZ_{k+0.5} = \frac{DS_{k+0.5} - \overline{DS}_*}{\overline{SD}}$$

APPENDIX 4

CALCULATION OF POOLED SMW DISTANCES

For each window midpoint location a pooled SMW dissimilarity estimate can be calculated from the individual Z score-transformed SMW dissimilarities from $s = 1, 2, 3, \dots, T$ different window widths by

$$\overline{DZ}_{k+0.5} = \frac{\sum_{s=1}^T DZ_{k+0.5,s}}{T}$$

APPENDIX 5

CALCULATION OF CONCENTRATION CONTOURS AROUND BIVARIATE GROUP CENTROIDS

Consider a bivariate, orthogonal set of $i = 1, 2, 3, \dots, N$ data points. Each point has ordinate and abscissa coordinates

of X_i and Y_i , respectively. The grand centroid of the data set has coordinate positions defined by

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} \text{ and } \bar{Y} = \frac{\sum_{i=1}^N Y_i}{N}.$$

Consider a new axis A , with angle α relative to the X axis and passing through the grand centroid of the set of data points. Each point i in the original XY coordinate system can be projected onto the new axis, with each coordinate location A_i being defined as relative to the grand centroid by

$$A_i = \cos \alpha (X_i - \bar{X}) + \sin \alpha (Y_i - \bar{Y}).$$

The new axis A has a mean (\bar{A}) of zero and standard deviation of

$$SD_\alpha = \frac{\left[\sum_{i=1}^N (A_i - \bar{A}_i)^2 \right]^{1/2}}{N - 1}.$$

Points at locations of ± 1 standard deviation from the grand centroid along axis A have positions in the original XY coordinate system of

$$X_{A\alpha} = \bar{X} \pm (\cos \alpha \cdot SD_\alpha) \text{ and } Y_{A\alpha} = \bar{Y} \pm (\sin \alpha \cdot SD_\alpha).$$

If the A axis is rotated at 5° intervals for 180° , and the calculations given above are repeated at each new axis position, a 1 standard deviation concentration contour is defined around the grand centroid of the set of points.