Quantitative methods and biogeographic analysis

Pierre Legendre **1990**

Département de sciences biologiques, Université de Montréal, C.P. 6128,
succursale Centre-ville, Montréal, Québec, Canada H3C 3J7

Pp. 9-34 *in*: Garbary, D. J. & G. R. South (eds.)
Evolutionary biogeography of the marine algae of the North Atlantic
NATO ASI Series, Vol. G 22. Springer-Verlag, Berlin
ISBN : 3-540-51761-8 and 0-387-51761-8
X + 429 pages

QUANTITATIVE METHODS AND BIOGEOGRAPHIC ANALYSIS

Pierre Legendre
Département de sciences biologiques
Université de Montréal
C.P. 6128, Succursale A
Montréal, Québec
Canada H3C 3J7

Abstract - Numerical methods are presented that represent three different approaches to biogeographic problems. The first approach is multivariate data analysis. The delineation of biogeographic "provinces" or areas is a type of descriptive analysis that can be accomplished by clustering faunal data (with or without spatial contiguity constraint) and drawing the resulting choropleth map. On the other hand, ecological biogeographers like to use ordinations of sampling localities and interpret the main axes of variation in terms of environmental gradients; canonical ordination, where a species presence or abundance data table and an environmental data matrix are both analyzed simultaneously, can be used with profit in this context. Secondly, the analysis of spatial patterns can help identify the type of spatial distribution of the biological material, both at the population and at the community level, while Mantel tests and other derived analyses make it possible to test hypotheses concerning causal factors possibly responsible for the observed spatial structures. Finally, phylogenetic-tree reconstruction methods, as well as other techniques, can be used for historical biogeographic studies; these include the study of taxa cladograms and of area cladograms.

INTRODUCTION

Myers & Giller (1988) have clearly demonstrated that historical and ecological biogeography can both contribute to explaining the distributions of organisms over the surface of the Earth. In both of these approaches, the hypotheses to be tested are developed by studying the non-randomness of the distribution patterns of living organisms in nature. While the distinction between a purely historical and a purely ecological perspective often depends on the spatial or temporal scale of the study, the two approaches must rely on sound numerical methods to support as efficiently as possible their respective hypotheses, using observational data.

This paper will focus on the most important numerical methods needed to demonstrate the non-randomness of the patterns found in nature and to approach biogeographic problems from an ecological and a historical perspective. [Asterisks (*) indicate specialized methods, not available in general-purpose statistical packages, for which computer programs are suggested in Table 5.]

NATO ASI Series, Vol. G 22
Evolutionary Biogeography of the
Marine Algae of the North Atlantic
Edited by D. J. Garbary and G. R. South

MULTIVARIATE DATA ANALYSIS

A large fraction of the currently published biogeographic studies concerns present-day faunas and floras. Typical problems involve the delineation of biogeographic "provinces" or areas, or the exploration of ecological relationships by correlating geographic distributions of organisms to environmental conditions. Methods of univariate or multivariate data analysis, coupled with classical geographic analysis, are the appropriate tools for such problems.

A quantitative study of a spatially structured phenomenon, like the geographic distribution of a species or a group of species, and of their controlling factors, should start with the study of distribution maps of the variables of interest: the species, and the environmental variables. Biogeographers usually satisfy themselves with rather unsophisticated kinds of maps. The two most common types are the following (both types can be produced by computer software).

1- Isoline maps are similar to the contoured maps used to represent altitudes on topographic maps, where the nested isolines represent different intensities of some continuous variable. Since these maps are in most cases derived from point samples obtained from the study area, intermediate values have to be estimated by interpolation. The sampling design is of prime importance for the accuracy of such reconstructions; this problem has been discussed by Burgess *et al.* (1981), McBratney *et al.* (1981) and Legendre *et al.* (1989), among others. Some of the common interpolation methods, including trend surface analysis* and kriging*, have recently been described for ecologists by Burrough (1987) and by Legendre & Fortin (1989). This type of mapping places the emphasis on the continuity properties of the variable to be mapped. The main trends in multivariate data can also be mapped by these methods, after reducing the data tables to synthetic variables, using ordination methods (examples below).

2- In so-called "choropleth maps" on the other hand, the study area is divided into nonoverlapping regions (or polygons), and boundaries are drawn around the regions on the basis that all points included in any given polygon allegedly possess some common property. This type of mapping is used to emphasize discontinuities in data. For multivariate data, the division of the area into regions can be obtained from clustering methods.

Clustering and ordination are two common tools of multivariate data analysis; they have been used by ecologists and biogeographers for a long time (Orlóci, 1978; Legendre & Legendre, 1983, 1984a; Jongman *et al.*, 1987; etc.) and are currently available in easy-to-use computer packages, both on mainframes and on microcomputers.

Clustering

The basic idea behind clustering is to partition a collection of objects into classes on the basis of their overall similarity or dissimilarity. (We will consider here only the clustering methods that produce nonoverlapping clusters, although methods exist that find overlapping clusters.) Many different clustering methods have been developed, based on mathematical models that are all adequate for producing discontinuous groups: geometric, graph theoretic, probabilistic, information theoretic, etc. Users also have to know that clustering can be obtained in two basically very different ways:

1- Many of the commonly used methods proceed in two steps: first an overall measure of resemblance (similarities or distances, reviewed in most textbooks of multivariate data analysis) is computed among all pairs of objects, which are sampling localities in most biogeographic studies; then clustering is performed, based upon this resemblance matrix. This family of methods includes linkage clustering (single, complete, intermediate); the average clustering methods (UPGMA, WPGMA, UPGMC, WPGMC), Ward's minimum-variance clustering, McQuitty's similarity analysis, flexible clustering, as well as the other methods that can be computed with the general agglomerative algorithm of Lance & Williams; and many of the divisive methods. "Eye-balling" methods, such as an ordination (principal components or correspondence analysis, for instance) followed by approximate delineation of clusters on the ordination diagram, can be included in this category. The choice of a resemblance measure appropriate to the data at hand is critical to obtaining meaningful results; guidance is provided by Legendre & Legendre (1983, 1984a) and by Gower & Legendre (1986).

2- In other methods, the assessment of similarity is done as part of the clustering method itself. These include the well-known *k*-means method, the ecologically-oriented method of information analysis (Williams *et al.*, 1966), as well as Lefkovitch's (1980, 1982, 1987) conditional clustering.

It is important to understand that clustering is a collection of heuristic procedures. This means that, in most cases, several methods and algorithms are equally appropriate, but none of them may also include assumptions related to the ecological or biogeographic theory of interest in the study. As a consequence, results also have to be judged heuristically. In practice, this means that among the methods that are appropriate to the data, users of clustering are encouraged to apply several methods to their data. Decision, as to the most interesting result among those that are technically appropriate, is often based on one of the following considerations, which are external to the clustering analyses:

- retain for interpretation only the groups that are evidenced by all methods used (strict consensus: see below);

- keep the partition or hierarchy that optimizes some objective criterion, such as the "cophenetic" or matrix correlation with the starting resemblance matrix;
- consider only the partition or hierarchy that best fits some other criterion of interest, such as geographic contiguity of the data points in each group (see also the methods of constrained clustering, below);
- consider only the partition or hierarchy that best fits some other data set of interest, for instance a matrix of environmental data (see also the methods of canonical ordination, below).

Let us remember that biogeographers often use clustering methods to delineate homogeneous geographic areas, based upon faunistic or floristic data. In this respect, clustering may be looked at as a family of methods — and indeed, as the only family of methods presently available — for mapping multivariate data. It is important at this point to realize that two different objectives may be pursued, and that they lead to two different alternatives in clustering.

In the first case, biogeographers are interested to locate all the areas where each given type of fauna or flora is found. Typically, one may want to test hypotheses of environmental control of species assemblages. In that case, geographically discontinuous ("disconnected") groups are just as interesting as continuous ("connected") areas, so that the results obtained by clustering areas on the basis of the species data can be mapped directly, and subsequently interpreted, using the environmental data.

On the other hand, the prime purpose of the study may be to outline connected geographic areas, based on the species data. Typical studies include the delineation of biogeographic "provinces", and the study of geographic dispersal of species assemblages after some event. In such studies, there is an implicit constraint that the geographic areas based on the species data will be of interest only if they are internally connected; that is, if none of them is split into geographically disconnected sub-areas. It may happen, of course, that the usual clustering methods will produce connected areas; an example is the study of Bock & Root (1981) who delineated biogeographic regions for landbird winter abundances in southern Canada and the United States. In other cases, however, some of the regions obtained by simple clustering may come out scattered over the study area. This is why several authors (Ray & Berry 1966; Webster & Burrough 1972; Lefkovitch 1978, 1980; Legendre & Legendre, 1984b; and others) have suggested including a constraint of spatial contiguity in the clustering algorithm itself*, in order to guarantee that the clustering results will also be geographically meaningful. Technically, this means that after the resemblance matrix has been computed among sites from the species data, one proceeds with any of the usual clustering methods, with the difference that the algorithm is constrained to cluster only those sites or site groups that are geographically contiguous. The constraint is provided to the program in the form of a list of connections, or spatial links, among neighbouring localities. Clustering with constraint has interesting properties. On the one hand, it reduces the set of mathematically possible solutions to those that are geographically meaningful. This avoids the well-known problem of clustering methods, where different solutions may obtain from different clustering algorithms applied to the same data set; constraining all these algorithms to produce results that are geographically consistent forces them to converge towards very similar solutions. On the other hand, the partitions obtained in this way reproduce a larger fraction of the structure's information than equivalent partitions obtained without constraint (Legendre, 1987). Finally, constrained agglomerative clustering is faster with large data sets than the unconstrained equivalent, because the search for "the next pair to join" is limited to adjacent groups only (Openshaw 1974; Lebart 1978). Examples of the use of constrained clustering for biogeographic analysis are found in Legendre & Legendre (1984b) on freshwater fishes, and in Legendre & Fortin (1989) on terrestrial vegetation.

Ordination

An ordination is a plot of object-points along an axis corresponding to a quantitative or (at least) an ordered relationship, or on a system of two or more such axes. Scatter diagrams, which are two-dimensional ordinations, are familiar to all scientists. The purpose of ordinations is to exhibit the salient features of a data set, such as gradients, groupings, or the presence of outliers. When the data contain more than two dimensions, instead of plotting all pairwise comparisons of variables, methods exist to find the "best" projection of the objects in one, two, or more dimensions; these graphs are called ordinations in a space of reduced dimensionality ("reduced space"). The best-known of these methods are principal components analysis, correspondence analysis, and principal coordinates analysis; several more specialized methods, often referred to as *scaling* methods, also exist; Gower (1987) gives a clear overview of these methods, and illustrates each one with vegetation data.

Ordination must be seen as the geometric problem of projecting points from a multidimensional space onto a smaller number of axes. The main ordination techniques differ from one another, mostly in the distance relationship they allow to preserve through that projection.

In *principal components analysis*, the axes on which the points are plotted in the original, multidimensional space are the original variables, so that the distance between any two points in that space can easily be calculated using the *Euclidean distance* formula. We say, then, that principal components analysis preserves the Euclidean distance, although distances are somewhat distorted when an n-dimensional data set is projected into 2 or 3 dimensions. Solutions of the so-called canonical equation give the various eigenvalues; to each one is associated an eigenvector, from which the principal components are computed. The principal components give the coordinates of the data points in the new space; that is, their positions on the successive principal axes. The method guarantees that the first principal axis passes through the direction of greatest variance of the scatter of points in the original, multidimensional space; the variance of the points along this first axis is given by its eigenvalue. Similarly, the succeeding principal axes, with variance equal to their respective eigenvalue, pass through the greatest successive directions of the n-dimensional scatter of points. The principal axes are orthogonal to one another, which means that their correlation to one another is zero. Since the first few principal axes account for the largest fractions of the data set variance, they are the most efficient axes to summarize, in a few important dimensions, the greatest part of the variability of a larger number of original variables.

*Correspondence analysis** differs from principal components analysis in the fact that the distance among points is the so-called *chi-square distance*, instead of the Euclidean distance. While the Euclidean distance is appropriate to represent relationships among points as measured, for instance, by environmental variables, or by other variables that can take positive, null, or negative values, the chi-square distance is appropriate to analyze either a species presence/absence or abundance table, or else a contingency table comparing two qualitative (categorical) variables (Legendre & Legendre, 1983, 1984a; Gower, 1987).

In *principal coordinates analysis**, as well as in the other forms of metric or non-metric scaling, the resemblance matrix is the user's choice, so that scientists can produce ordinations using any one of the metric or semimetric distance measures available in the literature (reviewed in most texts of numerical ecology), or from the corresponding similarity functions. In these ordination methods, the programs read in a previously computed resemblance matrix, and find the best projection of the data points in a space of reduced dimensionality. "Best" is defined by different criteria (depending on the method), the basic concept always being to distort, as little as possible, the original interpoint distances during the process of projecting the points onto the reduced space.

The chief role of ordinations in biogeography is to help identify gradients or other structures in the species abundance data obtained at the sampling localities, and to explain the spatial arrangement of the community structure using other environmental variables by *a posteriori* correlation or regression. This mode of analysis is called "indirect gradient analysis" (Whittaker, 1967; ter Braak, 1987b). The converse can also be done, where the ordination is done on the environmental variables, and the interpretation made using the species abundance data, either one at a time (single species biogeography) or all together (community biogeography). Reduced-space ordination of species abundance data is usually done by correspondence analysis or by principal coordinates analysis (or other form of scaling), while ordination of environmental variables is done either by principal components or by principal coordinates analysis (or other form of scaling). Ordinations can also be used in conjunction with clustering, and in particular with hierarchical clustering, to help decide what is the most informative partition (cutting level) in a dendrogram to produce a choropleth map. Finally, since ordination axes do represent new synthetic quantitative variables, they can be mapped directly (Fig. 1).

There are ways of jointly analyzing directly a species abundance data table and an environmental data matrix ("direct gradient analysis": ter Braak, 1987b), instead of relying on *a posteriori* correlations to associate environmental characteristics to species ordination axes. These methods are called canonical ordination (or constrained multivariate regression), and they pertain to two families: like principal components analysis, *canonical correlation analysis* and *redundancy analysis** preserve the Euclidean distance among points, while *canonical correspondence analysis** preserves the chi-square distance, just like standard correspondence analysis. I will briefly discuss the new method of canonical correspondence analysis (ter Braak, 1986, 1987a, 1987b) which was designed to interpret species assemblage variations by environmental variables, and should prove particularly useful for biogeographic analysis.

Direct and indirect gradient analyses will lead to the same results, if the major species gradients can readily be accounted for by the set of available environmental variables. It often

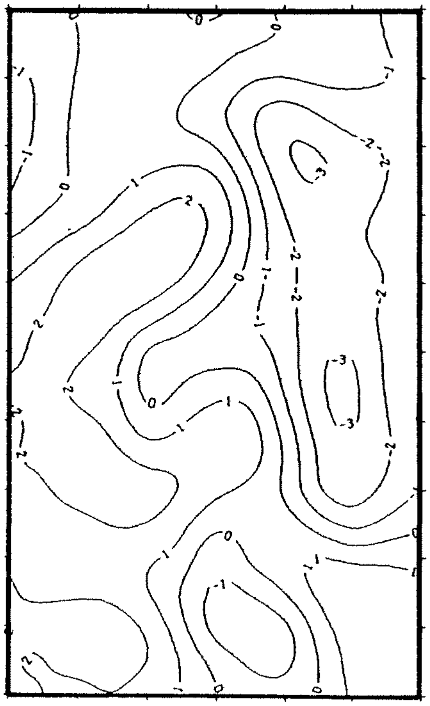

Fig. 1. Terrestrial vegetation data: 200 sampling stations, 28 tree species. Isoline map obtained by kriging from the first principal coordinate of a Steinhaus similarity matrix among sampling localities. Figure from Fortin (1985), with permission. The data set is also described in Legendre & Fortin (1989). The mapped area is 600 m wide and 1 km high. The negative areas of the map are the zones of highest density of the maple sugar *Acer saccharum*

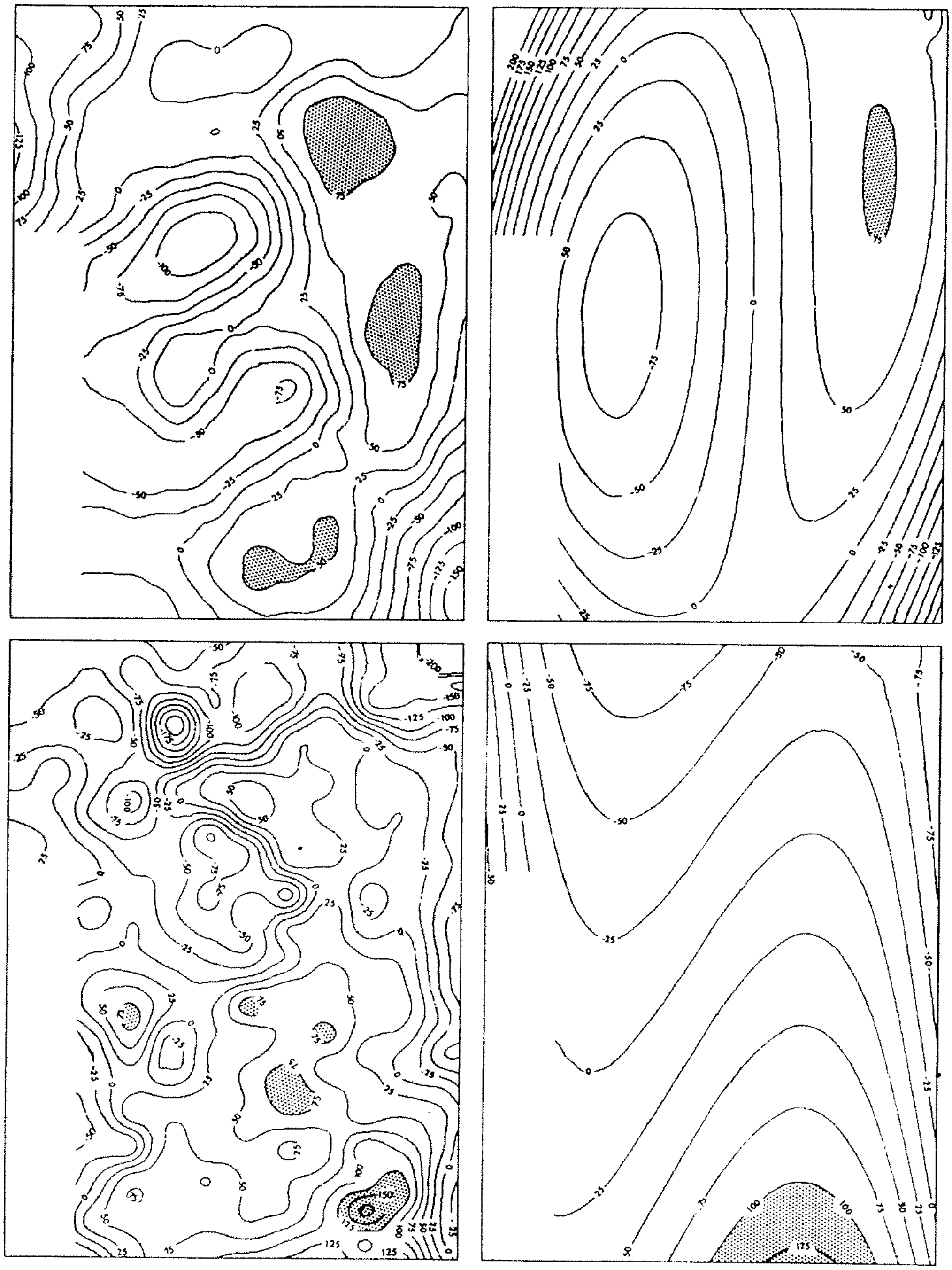

Fig. 2. Maps obtained by block kriging for the sample scores, on canonical axes I (top) and II (bottom), in the species space (left) and in the trend-surface geographic space (right); values multiplied by 100 for mapping. Peaks are shadowed. No samples had been taken from the blanked area on the left

happens, however, that a large part of the species variability is not directly controlled by environmental variables (bottom-up process), but, instead, by other factors, such as historical events, or by top-down ecological processes. In such cases, the larger part of the species variability, that occupies the first principal axes of the ordination, has nothing to do with environmental variables, so that their influence will not be brought out. Even though the environmental variables cannot explain a large part of the species variability, they may be useful to explain some of the remaining variation, and that fraction may still be important. The role of canonical ordination is to bring out that part of the species variability (first data table) that can be accounted for by the environmental variables (second data table). So, canonical ordination is related to multiple regression as well as to ordination. More precisely, the technique of canonical correspondence analysis selects as the first ordination axis the linear combination of environmental variables that maximizes the dispersion of the species scores; the species scores remain weighted averages of the site scores, as in the normal form of correspondence analysis, and the first eigenvalue measures the dispersion of the species scores along the axis. The second and further axes also select linear combinations of environmental variables that are maximally related to the species data, subject to the constraint that the successive fractions of the environmental variables' variability have to be uncorrelated with one another. Canonical correspondence analysis then produces a "constrained correspondence analysis" solution, since the site scores in the species space are still weighted averages of the species scores, which are also constrained, to be expressed as linear combinations of the environmental variables (ter Braak, 1987b).

In biogeographic analysis, the second data table can be made to contain the geographic coordinates (x, y) of the data points instead of environmental variables, as suggested by ter Braak (1987a). In that case, an ordination of the species data can be obtained that will be constrained to be consistent with the geographic distribution of the sampling localities. Columns can also be included in the geographic data table that contain values of x^2, y^2, xy, and possibly higher powers of the basic coordinates, to allow the species data to fit to a higher-degree polynomial of the basic geographic coordinates, in the manner of trend surface analysis (Burrough, 1987; Legendre & Fortin, 1989), which can be extended in this way to multivariate data. On the other hand, a biplot of species and geographic variables should indicate what species have the most important spatial structures. An example is presented for the 200 points of vegetation data of Figure 1. The geographic data matrix contains all the terms for a cubic trend surface regression of the form

$$z = b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3 \qquad (1)$$

Table 1 shows the weights of the species and of the geographic variables in the formation of the first two canonical axes. The first two eigenvalues are 0.23237 and 0.17912 respectively; they measure the amount of variation of the species data that is explained by the first and the second canonical axes and, hence, by the geographic variables; these first two canonical axes together account for 50% of the 28-species-geography biplot, and for 10% of the total 28-species variability, at the spatial scale determined by the sampling grid. Figure 2 shows maps obtained by kriging the sample scores: on the left, the sample scores are computed in the species space; and on the right in the geographic coordinates space. In other words, the sample data values from which the maps on the left-hand side have been interpolated are weighted averages of the species (weights given in Table 1), while the data values for the maps on the right-hand side are trend-surface maps based on equation (1) above, with weights also given in the Table. For canonical axis I (top portion of the Figure), the first community structure surface (left) is very

Table 1. Species and standardized geographic variable weights (canonical coefficients) for canonical axes I and II of the vegetation data study

Species	Axis I	Axis II	Geographic var.	Axis I	Axis II
1. *Abies balsamea*	-1.98	-1.55	1. x	-34.23	-20.02
2.*Acer pensylvanicum*	1.50	-3.17	2. y	-0.82	3.05
3. *Acer rubrum*	-0.22	0.93	3. x^2	71.88	44.81
4. *Acer saccharum*	1.29	0.43	4. xy	-5.29	-6.73
5. *Acer saccharinum*	-1.00	-3.87	5. y^2	5.76	-0.29
6. *Betula alleghaniensis*	-0.22	-1.14	6. x^3	-38.43	-24.73
7. *Betula papyrifera*	-0.97	-0.24	7. x^2y	6.91	3.34
8. *Betula populifolia*	-0.86	1.30	8. xy^2	-6.04	0.65
9. *Carpinus caroliniana*	-0.11	-1.21	9. y^3	0.41	-0.17
10. *Carya cordiformis*	0.96	-2.64			
11. *Fagus grandifolia*	2.66	-0.23			
12. *Fraxinus americana*	-0.86	-0.88			
13. *Fraxinus nigra*	-0.98	-1.72			
14. *Fraxinus pennsylvanica*	0.65	-2.45			
15. *Juglans cinerea*	0.65	0.17			
16. *Malus pumila*	-2.28	0.78			
17. *Ostrya virginiana*	1.11	-0.51			
18. *Pinus strobus*	-0.62	3.93			
19. *Populus balsamifera*	-2.10	2.28			
20. *Populus grandidentata*	2.18	-1.25			
21. *Populus tremuloides*	-1.07	-0.31			
22. *Prunus pensylvanica*	0.94	-0.64			
23. *Prunus serotina*	0.26	-0.60			
24. *Thuja occidentalis*	-2.58	0.09			
25. *Tilia americana*	1.13	-2.09			
26. *Tsuga canadensis*	-0.56	-0.75			
27. *Ulmus americana*	-1.35	-0.26			
28. *Ulmus rubra*	-2.07	0.17			

similar to the map in Figure 1, that represents the first principal coordinate obtained by unconstrained ordination of the vegetation data; the signs are inverted, a result of arbitrary decisions made by the computer programs and has no ecological significance. The species loadings in Table 1 show that the peaks of this map correspond to high densities of species 11, 20, 2 and 4, while the troughs correspond to high densities of species 24, 16, 19 and 28. The interesting fact is that this first component of the community structure can be pretty well approximated (correlation = 69%) by the trend surface map on the right, with high values in the upper right and the extreme upper left regions, and low values present in the center-left and in the lower right-hand regions; in other words, this equation predicts pretty well the community structure, from the simple knowledge of the geographic position of the samples. The same goes for axis II (lower part of Figure 2), where the second community structure surface (left) is well approximated by the trend surface on the right, to which it is correlated at 71%. This hints to a control of this community structure by mechanisms that make it strongly spatially structured. Leduc *et al.* (submitted MS) have investigated the hypothesis that while some of the tree species is this assemblage seem to be controlled by spatially autocorrelated environmental factors, other species' distributions are controlled by patch or gap dynamics.

Geographic variables can be used in yet another way by biogeographers: large-scale gradients in both the species and the environmental data can be removed by regression on the geographic coordinates (trend surface analysis). The study of the relations between species and environmental variables is then carried out on the regression residuals, either by direct or by indirect gradient analysis; this is an elegant way of removing this confounding factor, which is a common cause for spurious correlations in biogeographic analyses (Legendre & Troussellier, 1988). When removing geographic trends, however, one must always keep in mind that one may be discarding an important component of the information, if the community structure is indeed determined by a spatial process.

ANALYSIS OF SPATIAL PATTERNS

Much can be said on the importance of spatial heterogeneity in ecological and biogeographic theory. Several theories and models implicitly or explicitly assume that elements of an ecosystem that are close to one another in space are more likely to be influenced by the same generating processes. Organisms are not distributed evenly or at random in nature, for the good reason that spatial structures are functional in ecosystems, and not the result of some random process; a spatially homogeneous ecosystem simply could not function. This is why present-day biogeographers are interested in detecting the spatial arrangements of populations and communities. As mentioned in the previous section, mapping is a good way of grasping the salient features of geographic distributions. Qualitative examination of maps, when coupled with wishful thinking, may, however, play tricks on the best-intentioned of biogeographers. This is why it is important to be able to rely on some form of statistical analysis to confirm or infirm hypotheses about the observed spatial distributions of organisms and assemblages. This is the first role of spatial pattern analysis in biogeography. We will see that spatial analysis can go further and make it possible to test simple causal models that involve spatial variables as predictors.

This section will be brief, since the statistical analysis of spatial patterns for community analysis has been reviewed recently (Legendre & Fortin, 1989). I will limit the exposé to a guide to the most appropriate methods (with references), classified by the ecological and biogeographic questions they allow to be answered. Readers should refer to the above-mentioned paper for a more complete presentation of the methods, including formulas and ecological examples.

In the exploratory or hypothesis-generating phase of a research program, biogeographers should spend some effort testing for the statistical significance of the spatial structure they have detected or hypothesized. This can be done either in order to establish that there is a significant spatial structuring, and to determine the kind or shape of the pattern, or, to establish that there is no significant spatial pattern in the data, in order to validate the use of parametric statistical tests. Indeed, the presence of autocorrelation in data — spatial or temporal — invalidates the use of the classical tests, since the assumption of independence of the observations is not met. Testing procedures that are valid in the presence of spatial autocorrelation are found in Cliff & Ord (1981), Edgington (1987), Griffith (1987), Odland (1988) and Legendre *et al.* (1990).

There are various ways of testing for the presence of spatial autocorrelation in data:

1- A common structure encountered by biogeographers is the presence of a large-scale gradient in species composition. Such gradients may reflect a change in major environmental conditions, such as harshness of climate in north-south gradients, or dryness as one moves inland from an ocean margin. The Mantel test* (1967) of matrix correspondence is an easy and elegant way of verifying such an hypothesis. First, distance matrices are computed among the sampling localities, in two different ways. The distance matrix **X** is computed for the species composition data, using one of the many ecological resemblance coefficients adapted to species abundance data; an often used coefficient is the Odum (1950) or Bray & Curtis (1957) *percentage difference*. The second matrix **Y** contains the geographic distances among sampling localities, i.e., their distance in m or in km on the Earth surface. The null hypothesis to be tested is that the distances among points in matrix **X** are not linearly related to the corresponding distances in matrix **Y**; or in other words, that the species composition difference does not increase as the geographic distance among sampling stations increases. The normalized form of the Mantel statistic has the same formula as the Pearson product-moment correlation coefficient, computed for pairs of distance values in the off-diagonal of the two distance matrices. This statistic cannot be tested as a regular correlation coefficient, however, because distances in a matrix are not independent of one another; instead, it is tested either by a permutation procedure, or using a normal approximation which becomes asymptotically valid as the size of the matrices being compared increases. Schnell *et al.* (1986) have used the Mantel test to study patterns of geographic variation in dolphin morphology in the Pacific Ocean. The Mantel test has also been used to test the agreement of ecological data with a model (Burgman, 1987; Hudon & Lamarche, 1989; Legendre & Fortin, 1989), as well as hypotheses in panbiogeographic analysis (Page, 1987: see below).

2- In addition to gradients, it is possible to explore other shapes of spatial structures using *structure functions*. The most commonly used structure functions for univariate data (one species at a time) are the variogram* and the spatial correlograms*, and the Mantel correlogram* for multivariate data (whole species assemblages, as in the Mantel test). In particular, a *correlogram* is a graph where autocorrelation values are plotted in ordinate, against distances *d* among localities in abscissa. Figure 3 is an example of a correlogram for variable NO_3^- in the Thau marine lagoon, in southern France. In such an analysis, the problem is subdivided into a number of discrete distance classes, and an autocorrelation statistic is computed for each distance class. One must assume that a single dominant spatial structure exists over the study area; or, in other words, that the main large-scale structure is the same everywhere in that area. In correlograms, each statistic (each point in the correlogram) can be tested for statistical significance, and an overall test can be performed on the whole correlogram as well. Visual analysis of the shape of correlograms indicates whether the underlying spatial structure is aggregated, or has the form of a gradient, a wave, and so on. The correlogram in Figure 3 is characteristic of an aggregated structure; the second positive peak of the correlogram, at distance class 7 km, indicates the average distance between successive peaks or successive troughs of the variable, in the lagoon (see Legendre *et al.* (1989) for a description of

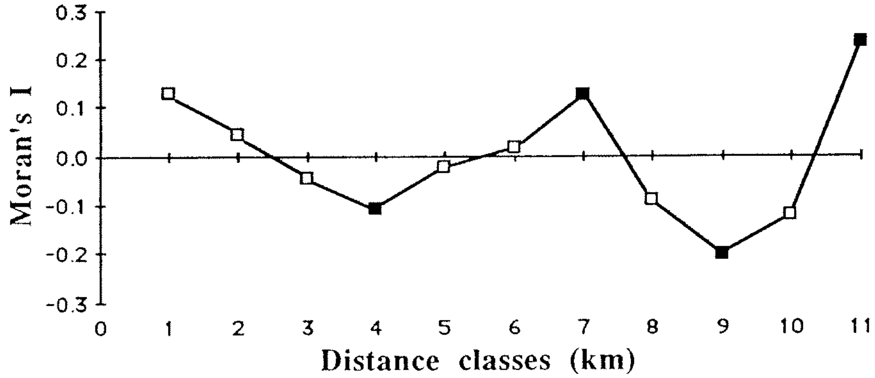

Fig. 3. Spatial correlogram for variable NO_3^- in the Thau lagoon on June 17, 1986. Black squares represent significant values of the Moran's I statistic at the 5% level; white squares are non-significant values

the lagoon and a mapping of this variable). Fig. 1 of Legendre & Fortin (1989) presents the correlograms of nine common types of spatial structures. The absence of a significant spatial structure, for the given observation scale, is indicated by a non-significant correlogram. For multivariate data, which in biogeography are often multi-species assemblages, the Mantel correlogram technique can be used (Sokal, 1986; Oden & Sokal, 1986); the end result is also a correlogram, where the statistics displayed in ordinate are Mantel statistics computed for the various distance classes, instead of Moran's *I* values. Structure functions allow biogeographers to describe the shape of their spatial structures at the same time as performing tests of statistical significance. In this way, they complement the visual exploration of data that can be done by mapping for univariate data, or by constrained clustering, and by normal or constrained ordination in the case of multivariate data (as explained in the previous section).

In the hypothesis-testing (model testing) phase of a research program, biogeographers often want to test models of assumed causal relationships among two or several variables or groups of variables. Legendre & Troussellier (1988) have shown that even highly significant correlations between biological and environmental variables can be spurious, because of the confounding effect of spatial autocorrelation. In other words, two variables or groups of variables may appear correlated simply because both obey the same spatial structure, which may be controlled by other factors that are not explicitly included in the model of assumed causal relationships, such as currents in the marine or estuarine environments. How is it possible to remove the effect of a common spatial structure from such studies of causal relationships? Smouse *et al.* (1986) have found a solution to this problem, suggesting that the variations in the two variables or groups of variables be expressed by two distance matrices, **A** and **B**, and that the effect of a third distance matrix **C**, that represents geographic relationships among sampling stations be removed, by a partial Mantel test procedure. Legendre & Troussellier (1988) have used this technique in an ecological modelling framework and have shown a case where the wrong ecological conclusion would have been reached using classical modelling techniques, had the confounding spatial structure not been studied.

I mentioned above that statistical tests adapted to the presence of spatial autocorrelation are becoming available. Among them, analysis of variance is of special interest to biogeographers, who are often interested in testing for significant differences among geographic areas on a map, for variables that are spatially autocorrelated. A statistically valid form of analysis of variance has been developed by Legendre *et al.* (1990) for these situations, where there is spatial autocorrelation in the variable being tested, as well as in the classification criterion, which is made of the geographic areas into which the map is subdivided. The problem with classical analysis of variance is that positive autocorrelation reduces within-group variability, thus artificially increasing the relative amount of among-group variance. Negative autocorrelation may produce the opposite effect. This difficulty can be viewed as a loss of an unknown number of degrees of freedom. Besides biogeography, such problems can be found in population genetics, in ecology and in other branches of biology, as well as in economics, epidemiology, geography, geology, marketing, political science, and sociology. A computer-intensive method has been developed to overcome this problem in certain cases. It is based on repeated permutations of the geographic areas on the map, while keeping them internally connected, followed by the computation of a pooled within-group sum of squares statistic after each permutation to produce a reference distribution to which the actual value, computed for the real map, can then be compared and its statistical significance assessed.

The paper of Legendre *et al.* (1990) presents a biogeographic example. The ecological hypothesis of interest was that the number of habitats available for freshwater fishes, in the James Bay territory of northwestern Québec (Canada), varied with the nature of the surface deposits; significant differences in the number of fish species were hypothesized among three contiguous geographic areas, dividing the territory under study according to geomorphology. A large-scale, well-known north-south gradient in fish numbers had first been eliminated by regression (trend surface analysis) from the 64, one-degree-square territory units, each about 7000 km^2 in area; this still left the data autocorrelated. Classical parametric ANOVA declared the differences among the three groups to be very highly significant ($p = 0.0009$). Since we know that spatial autocorrelation makes such a test too liberal — that is, differences among groups are too often declared significant when in fact they may not be — the analysis was repeated using the permutational procedure. After permuting 250 times at random the division of the territory into 3 regions, the probability of the null hypothesis was found to be 0.064, which is insufficient to reject the null hypothesis at the 0.05 level. Thus, we find no convincing evidence that fish species diversity, as measured by the number of species, varies in this territory according to the nature of the surface deposits. The modified test prevented a statistically erroneous decision from being made, which could have led to needless management decisions.

HISTORICAL BIOGEOGRAPHY

Historical (or phylogenetic) biogeography is interested in the simultaneous evolution of taxa, species assemblages (communities), and geographic areas. It is divided into a number of schools of thought, each one associated with the names of one or a few major proponents, without whom the vitality of the field would not be what it is now. I will briefly summarize these approaches, to set the stage for the presentation of the most commonly used numerical methods of data analysis, which is the aim of this paper.

The major division of the field is between *dispersal* and *vicariance* biogeography. The dispersal approach tries to follow the movements of organisms through geographic areas, while the vicariance approach studies how the creation of barriers, or other geologic processes, may have split up (vicariated) biota; leading to allopatric speciation. Vicariance is the property of closely related taxa that occupy similar biota located in distinct and often widely separated geographic areas. Dispersal and vicariance are probably constantly alternating processes in nature, although it may seem more appropriate to use the dispersal mode of analysis when analyzing intermediate time and space processes like postglacial reinvasion of a territory, for instance; on the contrary, the vicariance approach is certainly the main one of interest in studying long temporal and large spatial scale problems, like continental drift and its effects upon populations and communities, since it is at large scales that biogeographic patterns may involve related species groups that may have become separated by vicariance events (Brooks & McLennan, 1990). Important papers on vicariance biogeography are found in the two special issues recently devoted by *Systematic Zoology* to this question (Vol. 37 Nos. 3 and 4, 1988).

Vicariance biogeography is further divided into *panbiogeography* on the one hand, and *cladistic biogeography* on the other. Panbiogeography (Croizat 1952, 1958, 1981) uses "tracks" (lines on a map, each representing the distribution of one or more taxa), and "nodes" (regions with large numbers of tracks) to express phylogenetic relationships among areas. Tracks, represented by minimum-spanning trees, can in turn be translated into connectivity matrices, and Page (1987) has suggested using the Mantel test of matrix correspondence (see above) to assess the statistical significance of track concordance for different species.

Cladistic biogeography, on the other hand, uses cladograms as its main analytic tool. Cladograms can be used either to analyze the evolution of species (species cladograms, in phylogenetic systematics) or the biohistorical relationships of geographic areas (area cladograms). Among the methods of cladistic biogeography, Wiley (1988a) recognizes:

1- the reduced area cladogram of Rosen (1978, 1979), a method of analysis performed by hand, and now abandoned.

2- the component analysis of Platnick & Nelson (1978) and Nelson & Platnick (1981) — which has nothing to do with principal component analysis mentioned in the first section of this paper. Exposés of the method can be found in the above-mentioned papers, in Humphries *et al.* (1988), as well as in Wiley (1988a, b). Wiley finds major flaws in this method because it relies on consensus trees for the final solution. While they are good representations of the congruence of a set of cladograms, consensus trees may, when they contain unresolved polychotomies, turn out to be less parsimonious than the data allow when analyzed by parsimony methods.

3- Brooks *et al.* (1981) parsimony analysis of area relationships (below); and

4- Zandee & Roos (1987) compatibility analysis of area relationships (below), which is, in biogeographic analysis, another form of parsimony analysis and has nothing to do with the compatibility analysis used in phylogenetic analysis.

The relative merits of these methods have been thoroughly reviewed recently (Wiley, 1988a, b; Humphries *et al.*, 1988), and I will limit my exposé to a presentation of the principles of parsimony analysis in cladistic reconstruction, and then to a few statements about the way parsimony is applied to biogeographic analysis. Theriot (1989) reviews cladistic analysis in algal systematics.

Some basic principles of cladistic analyis are as follows: the first axiom of cladistic phylogenetics is that only derived states (called *apomorphic* states) contain information as to phylogenetic affinity (Mitchell 1901; Le Gros Clark & Sonntag 1926). If one considers a nominal variable observed over five species belonging to a common higher taxonomic category, then the first task of cladistic analysis is to formulate hypotheses as to the transformation series of the states of this character. Hennig (1950, 1966) suggests a number of methods for establishing the ancestral state of a character, as well as the transformation series; the best known, and most commonly applied, of these methods is outgroup analysis. In the case of a binary (two-state) character, the derived state, often called "1", identifies all the species that ought to be grouped together, under the assumptions that: a) the mutation that led to the derived state is likely to have occurred only once in the lineage, and, b) no reversal ("1" ⟶ "0") is likely to have occurred during the course of evolution. Thus, a single binary character leads to an incompletely resolved tree, in which all species sharing state "0" form a polytomy (or "bush") at the base of the tree, while higher up the tree, all species sharing state "1" form another polytomy that represents the monophyletic group. Consider now several binary characters: if the derived state is known in each case, then each derived state is the basis for recognizing a monophyletic group of species. If the characters are compatible, drawing all these monophyletic groups will lead to a more completely resolved tree. In the artificial example of Table 2, proposed by Felsenstein (1982), the first four characters are compatible and lead to a completely resolved tree for the five species (Fig. 4), with monophyletic groups **ac, ace, bd** and **d**. What happens now if we consider also characters 5 and 6? They show the existence of a new phylogenetic group made of species **b, c** and **e**, which is incompatible with the tree derived from characters 1 to 4. One can of course try to revise the ancestral state hypotheses; if the incompatibilities persist, two main solutions are offered:

1) allow parallel evolution (recurrent appearance of the derived state in various branches of the tree), as well as reversals towards the ancestral state ("1"⟶"0"), to occur within the reconstructed tree. In that case, one will be likely to find the tree that *minimizes* the number of such extra evolutionary events, that contradict the assumptions above. This is the principle of *parsimony*, or minimum-evolution in phylogenetic reconstruction (Edwards & Cavalli-Sforza, 1963, 1964); it states that the most likely phylogenetic reconstruction is that which minimizes the number of extra evolutionary events, due to reversals, parallelism, convergence or inadequate coding of characters.

2) identify the largest subset of characters that are compatible with one another, in the sense that they lead to a phylogenetic tree without extra evolutionary steps, and derive the cladogram using these characters only. This is the method of *compatibility* in phylogenetic reconstruction (LeQuesne, 1969; Estabrook, 1972; McMorris, 1975; etc.). It has nothing to do with the

biogeographic method of component compatibility briefly described below; phylogenetic trees used as the basis for cladistic biogeographic analysis may have been derived either by phylogenetic parsimony or by compatibility analysis.

A third solution, actually used in biogeographic analysis, consists of finding a consensus among the various single-character cladograms or the taxon-area cladograms. Notice that the very idea of consensus underlies a variety of different concepts (Adams (1972) consensus, strict consensus, majority rule, and so on. See the review paper by McMorris *et al.* (1982), as well as the special issue devoted by the *Journal of Classification* to the subject (Day, 1986)). In historical biogeographic analysis, this solution was tried in the context of component analysis and has been criticized by various authors, including Wiley (1988a) and Humphries *et al.* (1988), as mentioned above.

The parsimony method most often used in vicariance biogeography is to construct a Wagner tree (Kluge & Farris, 1969). Actually, the Wagner method first produces a network — that is, an undirected, or unrooted tree — that simply minimizes the number of evolutionary events, or, in other words, the number of mutations from "0" to "1" or from "1" to "0"; minimizing the total number of evolutionary events is actually equivalent to minimizing the

Table 2. Hypothetical binary data table

Taxa	Characters					
	1	2	3	4	5	6
a	1	1	0	0	0	0
b	0	0	1	0	1	1
c	1	1	0	0	1	1
d	0	0	1	1	0	0
e	0	1	0	0	1	1

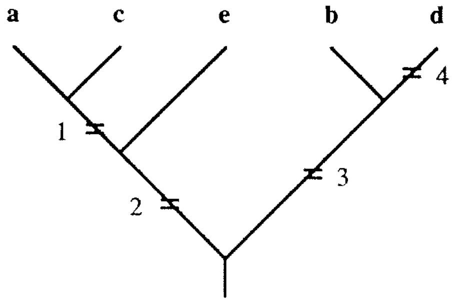

Fig. 4. Cladogram for the first 4 characters of Table 2

number of extra evolutionary steps. The tree is then rooted at the node that attaches the outgroup, real or hypothetical, that has been included in the analysis. The Wagner method, like many of the other methods of phylogenetic tree reconstruction (Table 3), was defined for binary characters. In phylogenetic analysis, multi-state characters whose transformation sequences are known, can be transformed into series of binary characters using the method of Kluge & Farris (1969), for instance. For quantitative characters, other methods have been proposed, based on the analysis of distance matrices (Fitch & Margoliash, 1967; Farris, 1972). See Felsenstein (1982) for a review of the phylogenetic-tree reconstruction methods, and Legendre (1986) for an application of several of these methods to community dispersal analysis from (species × area) data matrices (presence/absence, and frequencies of occurrence). Widely distributed computer packages to carry out these analyses are HENNIG 86*, MacCLADE*, PAUP* and PHYLIP*; Wagner trees can also be reconstructed using the CLADAREA* package. Among programs that perform parsimony analysis, and algorithms within the same program, differences may exist in their ability to find quickly the most parsimonious solution, or to find at all a solution that is close to the most parsimonious one; the reason is that the search for the most parsimonious solution represents a NP-complete problem, a computationally difficult problem whose solution can only be guaranteed by enumerating all possible trees; complete enumeration rapidly becomes impossible as the number of OTUs increases.

Table 3. Differences among the main methods of phylogenetic tree reconstruction for binary characters. See Felsenstein (1982) for details

Method	Minimize	No more than once per character	Forbidden
Parsimony methods:			
Wagner	$0 \rightarrow 1$, $1 \rightarrow 0$	-----	-----
Camin-Sokal	$0 \rightarrow 1$	-----	$1 \rightarrow 0$
Dollo	$1 \rightarrow 0$	$0 \rightarrow 1$	(01)[1]
Polymorphism	(01)[1]	$0 \rightarrow (01)$[1]	$1 \rightarrow 0$
Compatibility:	eliminate incompatible characters	-----	$0 \rightarrow 1 \rightarrow 0$ $1 \rightarrow 0 \rightarrow 1$

---

[1] (01) represents the polymorphic state. The polymorphism parsimony method minimizes the duration of retention of that state.

Table 5. The following programs are available to compute the various methods of biogeographic analysis identified by asterisks (*) in the text. Only those methods are listed that are not available in general-purpose statistical packages. Footnotes indicate addresses for programs

PROGRAM	METHODS
CANOCO[1]	Simple and canonical correspondence analysis, redundancy analysis.
CLADAREA[2]	Area cladistic analysis. Wagner tree reconstruction (parsimony).
COMPONENT[3]	Construction and comparison of area cladograms.
CORR2D[4]	Two-dimensional correlogram.
GEO-EAS[5]	Variogram, kriging.
GEOLIN[6]	Variogram, kriging. Also GEOSREC for non-linear and BLUEPACK for non-stationary kriging.
GEOSTAT[7]	Variogram, kriging.
HENNIG 86[8]	Wagner tree reconstruction (parsimony).
Kellogg's[9]	Variogram, kriging.
MacCLADE[10]	Interactive analysis of phylogenies and character evolution (parsimony).
NTSYS-PC[11]	Simple Mantel test; correspondence analysis.
PAUP[12]	Phylogenetic tree reconstruction (parsimony).
PHYLIP[13]	Phylogenetic tree reconstruction (parsimony, compatibility).
"R"[14]	Spatial autocorrelation, simple and partial Mantel tests, Mantel correlogram, clustering with space and time contiguity constraint. Connecting networks.
SAAP[15]	Spatial autocorrelograms (Moran's I and Geary's c).
SASP[16]	Two-dimensional spectral analysis.
SURFER[17]	Kriging from linear variogram only; other interpolation methods.
SYMAP[18]	Trend surface analysis; other interpolation methods.
UNIMAP[19]	Variogram, kriging; other interpolation methods.

[1] Cajo J. F. ter Braak, Agricultural Mathematics Group, TNO Institute for Applied Computer Science, Box 100, NL-6700 AC Wageningen, Netherlands.

[2] Willem N. Ellis, Instituut voor Taxonomische Zoologie, afd. Entomologie, Plantage Middenlaan 64, NL-1018 DH Amsterdam, Netherlands.

[3] Roderic D. M. Page, Department of Zoology, University of Auckland, Private Bag, Auckland, New Zealand.

[4] Applied Biostatistics Inc., 100 North Country Road, Bldg. B, Setauket, New York 11733, USA. (written by G. M. Jacquez)

[5] Evan J. Englund (Geo-EAS), USEPA EMSL-LV, EAD, P.O. Box 93478, Las Vegas, Nevada 89193-3478, USA. (developed by EPA)

[6] Geovariances International, 1 rue Charles-Meunier, F-77210 Avon-Fontainebleau, France.

[7] Geostat Systems International Inc., 4385 rue Saint-Hubert, Suite 1, Montréal, Québec, Canada H2J 2X1.

[8] James S. Farris, 41 Admiral Street, Port Jefferson Station, New York, NY 11776, USA.

[9] Computer Laboratory, W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, Michigan 49060, USA.

Table 5. cont.

[10] Wayne P. Maddison and David R. Maddison, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA.

[11] PC version from Applied Biostatistics Inc., 100 North Country Road, Bldg. B, Setauket, New York 11733, USA. (developed by F.J. Rohlf)

[12] David L. Swofford, Illinois Natural History Survey, University of Illinois, Urbana, Illinois 61801, USA.

[13] Joseph Felsenstein, Department of Genetics, SK-50, University of Washington, Seattle, Washington 98125, USA.

[14] developed by Alain Vaudor (P. Legendre's lab: see title page), is available for Macintosh microcomputers, VAX, and IBM mainframes.

[15] Daniel Wartenberg, Department of Environmental and Community Medicine, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, New Jersey 08854, USA.

[16] E. Renshaw, Department of Statistics, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom.

[17] Golden Software Inc., P. O. Box 281, Golden, Colorado 80402, USA

[18] No longer distributed by Laboratory for Computer Graphics and Spatial Analysis, Harvard University, USA. It is however still available at many computing centers.

[19] European Software Contractors A/S, Nørregade, DK-2800 Lyngby, Denmark.

relating distributional and environmental data through ordination; the statistical analysis of spatial patterns, and biohistorical reconstruction.

This presentation of numerical methods would be of little utility to field biogeographers if they were not given access to computer software capable of carrying out these calculations. Accordingly, computer programs, for methods that are not readily available in general-purpose statistical packages, are listed in Table 5.

ACKNOWLEDGEMENTS

This is publication No. 369 from the Groupe d'Écologie des Eaux douces, Université de Montréal. I am indebted to Pierre Drapeau, Ph.D. student at Université de Montréal, who directed the sampling program that produced the data used in Figures 1 and 2. Fruitful discussions with Dr. Cajo J.F. ter Braak during the international symposium on *Progress in theoretical vegetation science* held in Vienna in July 1988 led to the formulation of the application of canonical correspondence analysis to produce a space-constrained ordination, as illustrated in Figure 2. This study was supported by NSERC grant No. A7738 to P. Legendre.

REFERENCES

Adams EN III (1972) Consensus techniques and the comparison of taxonomic trees. Syst Zool 21: 390-397

Bock CE, Root TL (1981) Winter abundance patterns of landbirds in the United States and southern Canada. Amer Birds 35: 891-897

Bray RJ, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27: 325-349

Brooks DR, Thorson TB, Mayes MA (1981) Freshwater stingrays (Potamotrygonidae) and their helminth parasites; testing hypotheses of evolution and coevolution. pp. 147-175. In: Funk VA, Brooks DR (eds) Advances in cladistics, 1. Columbia Univ. Press, New York

Brooks DR, McLennan DA (1990) Historical ecology as a research program. In: Mayden RL (ed) Historical ecology and phylogeny of North American freshwater fishes. Stanford Univ. Press, Stanford

Burgess TM, Webster R, McBratney AB (1981) Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. J Soil Sci 32: 643-659

Burgman MA (1987) An analysis of the distribution of plants on granite outcrops in southern Western Australia using Mantel tests. Vegetatio 71: 79-86

Burrough PA (1987) Spatial aspects of ecological data. Chapter 7, pp. 213-251. In: Jongman RHG, ter Braak CJF, van Tongeren OFR (eds) Data analysis in community and landscape ecology. Pudoc, Wageningen

Cliff AD, Ord JK (1981) Spatial processes: models and applications. Pion Limited, London

Croizat L (1952) Manual of phytogeography. Junk, The Hague

Croizat L (1958) Panbiogeography. Published by the author, Caracas, Venezuela

Croizat L (1981) Biogeography: past, present, and future. pp. 501-523. In: Nelson G, Rosen DE (eds) Vicariance biogeography, a critique. Columbia Univ. Press, New York

Day WHE (1986) Foreword: comparison and consensus of classifications. J Classif 3: 183-185

Edgington ES (1987) Randomization tests, 2nd ed. Marcel Dekker Inc., New York

Edwards AWF, Cavalli-Sforza LL (1963) The reconstruction of evolution. Ann Hum Genet 27: 105

Edwards AWF, Cavalli-Sforza LL (1964) Reconstruction of evolutionary trees. pp. 67-76. In: Heywood VH, McNeill J (eds) Phenetic and phylogenetic classification. Systematic Association Publ No. 6, London

Estabrook GF (1972) Cladistic methodology: a discussion of the theoretical basis for the induction of evolutionary history. Ann Rev Ecol Syst 3: 427-456

Farris JS (1972) Estimating phylogenetic trees from distance matrices. Amer Nat 106: 645-668

Felsenstein J (1982) Numerical methods for inferring evolutionary trees. Quart Rev Biol 57: 379-404

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155: 279-284

Fortin M-J (1985) Analyse spatiale de la répartition des phénomènes écologiques: méthodes d'analyse spatiale, théorie de l'échantillonnage. Mémoire de Maîtrise, Université de Montréal

Gower JC (1987) Introduction to ordination techniques. pp. 3-64. In: Legendre P, Legendre L (eds) Developments in Numerical Ecology. NATO ASI Series, Vol. G 14. Springer-Verlag, Berlin

Gower JC, Legendre P (1986) Metric and Euclidean properties of dissimilarity coefficients. J Classif 3: 5-48

Griffith DA (1987) Spatial autocorrelation — A primer. Resource Publications in Geography, Association of American Geographers, Washington, D.C.

Hennig W (1950) Grundzüge einer Theorie der phylogenetischen Systematik. Deutscher Zentralverlag, Berlin

Hennig W (1966) Phylogenetic systematics. University of Illinois Press, Urbana, Illinois

Hudon C, Lamarche G (1989) Niche segregation between American lobster *Homarus americanus* and rock crab *Cancer irroratus*. Mar Ecol Prog Ser 52: 155-168

Humphries CJ, Ladiges PY, Roos M, Zandee M (1988) Cladistic biogeography. pp. 371-404. In: Myers AA, Giller PS (eds) Analytical biogeography — An integrated approach to the study of animal and plant distributions. Chapman and Hall, London

Jongman RHG, ter Braak CJF, van Tongeren OFR (eds) (1987) Data analysis in community and landscape ecology. Pudoc, Wageningen

Kluge AG, Farris JS (1969) Quantitative phyletics and the evolution of anurans. Syst Zool 18: 1-32

Lapointe F-J, Legendre P — A statistical framework to test the consensus of two nested classifications. Syst Zool 39: 1-14

Lebart L (1978) Programme d'agrégation avec contraintes (C.A.H. contiguïté). Cah Anal Données 3: 275-287

Leduc A, Drapeau P, Bergeron Y, Legendre P (submitted) Importance of spatial phenomena in the explanation of forest cover variations.

Le Gros Clark WE, Sonntag CF (1926) A monograph of *Orycteropus afer*. III. The skull. Proc Gen Meet Sci Bus, Zool Soc London 1926: 445-485

Lefkovitch LP (1978) Cluster generation and grouping using mathematical programming. Math Biosci 41: 91-110

Lefkovitch LP (1980) Conditional clustering. Biometrics 36: 43-58

Lefkovitch LP (1982) Conditional clusters, musters, and probability. Math Biosci 60: 207-234

Lefkovitch LP (1987) Species associations and conditional clustering: clustering with or without pairwise resemblances. pp. 309-331. In: Legendre P, Legendre L (eds) Developments in Numerical Ecology. NATO ASI Series, Vol. G 14. Springer-Verlag, Berlin

Legendre L, Legendre P (1983) Numerical ecology. Developments in environmental modelling, 3. Elsevier, Amsterdam

Legendre L, Legendre P (1984a) Écologie numérique, 2ième éd. Tome 2: La structure des données écologiques. Masson, Paris et les Presses de l'Université du Québec

Legendre P (1986) Reconstructing biogeographic history using phylogenetic-tree analysis of community structure. Syst Zool 35: 68-80

Legendre P (1987) Constrained clustering. pp. 289-307. In: Legendre P, Legendre L (eds) Developments in Numerical Ecology. NATO ASI Series, Vol. G 14. Springer-Verlag, Berlin

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

Legendre P, Legendre V (1984b) Postglacial dispersal of freshwater fishes in the Québec peninsula. Can J Fish Aquat Sci 41: 1781-1802

Legendre P, Oden NL, Sokal RR, Vaudor A, Kim J (1990) Approximate analysis of variance of spatially autocorrelated regional data. J Classif 7: 53-75

Legendre P, Troussellier M (1988) Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. Limnol Oceanogr 33: 1055-1067

Legendre P, Troussellier M, Jarry V, Fortin M-J (1989) Design for simultaneous sampling of ecological variables: from concepts to numerical solutions. Oikos 55: 30-42

LeQuesne WJ (1969) A method of selection of characters in numerical taxonomy. Syst Zool 18: 201-205

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res 27: 209-220.

McBratney AB, Webster R, Burgess TM (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. Computers & Geosciences 7: 331-334

McMorris FR (1975) Compatibility criteria for cladistic and qualitative taxonomic characters. pp. 399-415. In: Estabrook GF (ed) Proceedings of the Eight International Conference on Numerical Taxonomy. Freeman, San Francisco

McMorris FR, Meronk DB, Neumann DA (1982) A view of some consensus methods for trees. pp. 122-126. In: Felsenstein J (ed) Numerical taxonomy. NATO ASI Series, Vol. G 1. Springer-Verlag, Berlin

Mickevich MF (1981) Quantitative phylogenetic biogeography. pp. 209-222. In: Funk VA, Brooks DR (eds) Advances in cladistics, 1. Columbia Univ. Press, New York

Mitchell PG (1901) On the intestinal tract of birds, with remarks on the valuation and nomenclature of zoological characters. Trans Linnean Soc London (Zool Ser 2) 8: 173-275

Myers AA, Giller PS (1988) Process, pattern and scale in biogeography. pp. 3-12. In: Myers AA, Giller PS (eds) Analytical biogeography — An integrated approach to the study of animal and plant distributions. Chapman and Hall, London

Nelson G, Platnick NI (1981) Systematics and biogeography: cladistics and vicariance. Columbia Univ. Press, New York

Oden NL, Sokal RR (1986) Directional autocorrelation: an extension of spatial correlograms to two dimensions. Syst Zool 35: 608-617

Odland J (1988) Spatial autocorrelation. Scientific Geography Series, Vol. 9. Sage Publications Inc., Newbury Park, California

Odum EP (1950) bird populations of the Highlands (North Carolina) plateau in relation to plant succession and avian invasion. Ecology 31: 587-605

Openshaw S (1974) A regionalisation program for large data sets. Computer Appl 3-4: 136-160

Orlóci L (1978) Multivariate analysis in vegetation research, 2nd ed. Dr. W. Junk B.V., The Hague

Page RDM (1987) Graphs and generalized tracks: quantifying Croizat's panbiogeography. Syst Zool 36: 1-17

Page RDM (1988) Quantitative cladistic biogeography: constructing and comparing area cladograms. Syst Zool 37: 254-270

Platnick NI, Nelson G (1978) A method of analysis for historical biogeography. Syst Zool 27: 1-16

Ray DM, Berry BJL (1966) Multivariate socioeconomic regionalization: a pilot study in central Canada. pp. 75-130. In: Ostry S, Rymes T (eds) Papers on Regional Statistical Studies. Univ. of Toronto Press, Toronto

Rosen DE (1978) Vicariant patterns and historical explanation in biogeography. Syst Zool 27: 159-188

Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative geography. Bull Am Mus Nat Hist 162: 267-376

Schnell GD, Douglas ME, Hough DJ (1986) Geographic patterns of variation in offshore spotted dolphins (*Stenella attenuata*) of the eastern tropical Pacific Ocean. Marine Mammal Science 2: 186-213

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst Zool 35: 627-632

Sokal RR (1986) Spatial data analysis and historical processes. pp. 29-43. In: Diday E et al. (eds) Data analysis and informatics, IV. Proceedings of the Fourth International Symposium on Data Analysis and Informatics, Versailles, France, 1985. North-Holland, Amsterdam

ter Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167-1179

ter Braak CJF (1987a) The analysis of vegetation-environment relationships by canonical correspondence analysis. Vegetatio 69: 69-77

ter Braak CJF (1987b) Ordination. Chapter 5, pp. 91-173. In: Jongman RHG, ter Braak CJF, van Tongeren OFR (eds) Data analysis in community and landscape ecology. Pudoc, Wageningen

Theriot E (1989) Phylogenetic systematics for phycology. J Phycol 25: 407-411

Webster R, Burrough PA (1972) Computer-based soil mapping of small areas from sample data. I. Multivariate classification and ordination. II. Classification smoothing. J Soil Sci 23: 210-221, 222-234

Whittaker RH (1967) Gradient analysis of vegetation. Biol Rev 49: 207-264

Wiley EO (1988a) Vicariance biogeography. Ann Rev Ecol Syst 19: 513-542

Wiley EO (1988b) Parsimony analysis and vicariance biogeography. Syst Zool 37: 271-290

Williams WT, Lambert JM, Lance GN (1966) Multivariate methods in plant ecology. V. Similarity analyses and information-analysis. J Ecol 54: 427-445

Zandee M, Roos MC (1987) Component-compatibility in historical biogeography. Cladistics 3: 305-332