Advances in

ECOLOGICAL RESEARCH

Edited by

A. MACFADYEN

Department of Biology, University of Ulster,
Coleraine, County Londonderry, Northern Ireland

E. D. FORD

Institute of Terrestrial Ecology, Bush Estate,
Penicuik, Midlothian, Scotland

VOLUME 14

1984

ACADEMIC PRESS

(Harcourt Brace Jovanovich, Publishers)
London Orlando San Diego New York
Toronto Montreal Sydney Tokyo

ACADEMIC PRESS, INC. (LONDON) LTD.
24-28 Oval Road,
London NW1 7DX

United States Edition published by
ACADEMIC PRESS, INC.
Orlando, Florida 32887

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 62-21479

ISBN 0-12-013914-6

PRINTED IN THE UNITED STATES OF AMERICA

84 85 86 87 9 8 7 6 5 4 3 2 1

Bray–Curtis Ordination: An Effective Strategy for Analysis of Multivariate Ecological Data

EDWARD W. BEALS

I. INTRODUCTION

Since Goodall (1954) first applied an objective mathematical technique to the ordering of plant communities, techniques of ordination have proliferated to the point of overwhelming even the most mathematical of ecologists. Comparisons of various techniques began with Austin and Orloci (1966) and Bannister (1968), and have continued to be published in some quantity.

The ordination method of Bray and Curtis (BC) (1957), one of the few techniques specifically designed for phytosociological data, has stood up well in most comparisons (specific studies are cited later). In teaching a course entitled "Community Analysis" to graduate students at the University of Wisconsin, I have required of them a report comparing multivariate methods on a data set of their own choosing. The results of those

ISBN 0-12-013914-6

students who used phytosociological data, analogous zoological data, or resource-partitioning data also suggest that BC ordination is one of the better methods.

This article arises from two concerns. One is a general tenor in current literature that BC ordination is outmoded or certainly inferior to other methods (namely, reciprocal averaging or principal component analysis). The other is that outside Wisconsin, BC ordination is still limited mostly to the specific techniques Bray and Curtis (1957) originally proposed or the few modifications suggested by Beals (1960). I became aware of this when we obtained Cornell University's ORDIFLEX package of computer programs (Gauch, 1977). Several modifications of the original method are now such common practice at Wisconsin that we take them for granted. Although various papers have presented these improvements in the literature (Beals, 1965a,b, 1969a, 1973) and they have been used in other studies (such as Emlen, 1972, 1977; Lechowicz and Adams, 1974; Stephenson, 1974; Will-Wolf, 1980), they perhaps need to be presented as a formal body of options designed to improve the performance of the Bray–Curtis method.

Over many years I and my students have experimented with ordination methods and accumulated much information and many insights. A general survey of the comparative results of our studies is given in Section VI. Table 1 gives a summary of the methods which were tested in these various studies and which are discussed at length in this article.

II. ECOLOGICAL SPACES

Ordination implies an abstract space in which entities form a constellation. Each entity is located in that space on the basis of a set of attributes. A primary assumption is that the points are not located randomly but that there are correlations (in the broad sense) among the attributes. The object of ordination is to find major axes of variation through this constellation, to reduce the many dimensions of the system to a very few, with minimum loss of information. Ordination is a projection of a multidimensional system (not necessarily even Euclidean) onto a two- or three-dimensional map. This reduction is analogous to the process of classification, in which many entities are reduced to a relatively few categories. In both cases, the analysis should enhance the clarity of major patterns of variation, but it will obscure minor variation, which is presumably less important or even random.

Traditionally, beginning with the ordinations of Goodall (1954) and Bray and Curtis (1957), the entities are samples (stands) and the attributes are species values in those samples. This is vegetation space (if the species are plants), and the species in some sense represent dimensions. A more general

Table 1

Summary of the Most Frequently Cited Ordination Methods in This Article

Abbreviation	Method	Primary reference[a]	Feature maximized	Maximization strategy[b]
BC	Bray–Curtis sample reference points[c]			
—	Maximum range	Bray and Curtis (1957)	Range	A
—	Variance regression	This article	Linear gradient of point clusters	A
	Synthetic reference points			
AA–BC	Association analysis	This article	Interspecific heterogeneity	O
SS–BC	Sums of squares	This article	Variance between groups	O
RA	Reciprocal averaging	Hill (1973)	Correspondence of species and samples	A
DCA	Detrended correspondence analysis[c]	Hill and Gauch (1980)	Unknown	A
PCA	Principal components analysis	Orloci (1966)	Variance among samples	A
Kruskal	Global nonmetric scaling	Kruskal (1964)	Monotonicity between ordination and data distances	O

[a]Primary references are those most appropriate for a description of methodology, usually in an ecological context.
[b]A, Maximization along one axis at a time; O, maximization along all axes of ordination at once.
[c]Methods developed specifically for ecological data.

term, to include study of animal communities (Whittaker, 1952; Beals, 1960), is sociological space, since the system is analyzed at the community level (cf. phytosociology). Samples may be on any scale, from small individual quadrats (Beals, 1965a) to widespread community types (Curtis, 1959).

Previously (Beals, 1973), I suggested that the idea of species as dimensions of vegetation space should be abandoned in favor of a scalar approach, the concept of Δ-vegetation space. This is a space reflected simply by change in vegetation from point to point without any reference to individual species. More recently, at meetings, I have heard it suggested that the attempt to understand the spatial concepts is irrelevant and even counterproductive to the goal of gradient analysis. Nevertheless, there are spatial assumptions underlying all ordination. The dimensionality of the system represented by the data matrix is a direct function of the number of attributes (species) whether one considers those species as Euclidean orthogonal axes or not. While much of the discussion perhaps has been counterproductive, understanding the spatial concepts underlying ordination is not irrelevant.

Species are not the only attributes of samples that can be measured. Structural or functional features of the biota, rather than taxonomic features, have been used (Knight, 1965; Knight and Loucks, 1969) to produce another spatial system for ordination. An additional set of attributes of these samples is that of their environmental characteristics, which produce a different spatial continuum. Environmental ordinations may be more or less direct, using major environmental factors as axes (Whittaker, 1952, 1956; Mowbray and Oosting, 1968), or indirect, using various ordination techniques to reduce dimensionality (Loucks, 1962; Austin, 1968; Mohler, 1981). Natural but compound environmental gradients such as elevation are commonly used (Beals, 1969b; Terborgh, 1971).

Whittaker (1967) called ordination by environmental factors "direct gradient analysis," and that by sociological factors "indirect gradient analysis," but these terms could be reversed with equal justification, if the interest is in species patterns as well as environmental patterns. The terms environmental ordination and sociological ordination are more descriptive. In any case, each sample has at least two sets of attributes, species and environmental factors, and we can view that sample as a point in both kinds of multidimensional space. In each space, the samples collectively produce a unique constellation of points.

The disadvantage of environmental ordination is that one must prejudge which are the important environmental factors to the vegetation or to the fauna. An environmental ordination may omit important variables; it is often biased toward those factors most easily measured; measured variables may be scaled wrong; and biotic patterns imposed by competition, predation, and other interactions are ignored. For example, along a rainfall gradient, the axis will show the same distance between samples with 20- and 40-cm annual rainfall as between those with 100 and 120 cm, whereas the vegetation might change more radically between the former samples than between the latter. The vegetation says that the former distance is ecologically greater than the latter. Biotic interactions may cause discontinuities in vegetation change along a continuous environmental gradient (Beals, 1969b), and such patterns would not show in the point scatter in an environmental ordination.

It is true that ecologists expect a sociological ordination to reflect the environmental space as the community responds to it. We overlay environmental factors on the ordination and detect environmental patterns, and our evaluation of an ordination is often dependent on the clarity of those patterns. However, the clarity of species is a more appropriate index to the success of a sociological ordination, because the gradients are compositional in nature, not environmental. Species differences between two samples do reflect their environmental differences, but in a highly integrated fashion, which includes differences in biotic interactions and in historical events. The environmental differences are automatically scaled according to overall species response. Therefore the ordination with the clearest species patterns reflects the environmental space the way the biotic community interprets it.

In contrast to the above sociological-environmental system of samples, we can view ecological systems somewhat differently. Species can be considered as entities and their success at different sites (samples) as attributes of the species. Thus we can transpose the data matrix and do an ordination of species (Beals, 1965b; Gittins, 1965) in attribute space. If we view the species as entities, we can expand the domain of attributes of those species to include not only the sites of success (habitat or microhabitat), but also the resources used, behavioral activities, morphological or physiological adaptations to the environment, etc. Ordination of species by their habitat or microhabitat preference, by their resource usage, or by their foraging behavior I call "niche ordination," in distinction from sociological ordination. Niche ordinations show how species in a community or set of communities partition their resources and/or their habitats.

Ordination of animal species by habitat or microhabitat has become a popular strategy (James, 1971; Green, 1971; Emlen, 1972, 1977; Conner and Adkinsson, 1977; Dueser and Shugart, 1979; Carey *et al.*, 1980; Sabo, 1980; Noon, 1981; and many others). Resources used have been employed as attributes in species ordinations (Wolf, 1975; Hanski and Koskela, 1977; Futuyma and Gould, 1979) as have foraging behaviors (Karr and James, 1975; Holmes *et al.*, 1979). Morphological attributes of species have been used not only in numerical taxonomy, but also in an ecological context (Karr and James, 1975; Ricklefs and Travis, 1980).

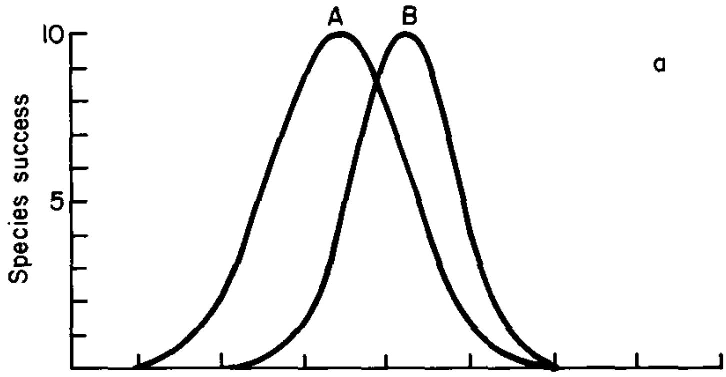
Niche and habitat data have been subject more commonly to numerical classification than to ordination, and even many of the above papers rely more heavily on their classification analysis, but ordination has demonstrated its value. Except for the studies emanating from Wisconsin (Emlen, 1972, 1977; Wolf, 1975, and several class projects), niche ordinations have been derived by various eigenvector techniques, most commonly principal component analysis. Bray–Curtis ordination was used successfully by Emlen and was compared with a variety of other techniques by Wolf. She found it to be clearly superior to principal component analysis. Therefore, zoologists analyzing niche space ought seriously to consider the Bray–Curtis approach.

This article is concerned with three types of ecological spaces: samples as points in compositional or sociological space, samples as points in environmental space, and species (or other taxa) as points in niche space (resource or habitat). Emphasis here is on the sociological space and to a lesser extent the niche space. As we shall see, these spaces may be better modeled as non-Euclidean spaces, although eventually we must project them into a Euclidean representation, that is, a graph of two or three dimensions.

III. BASIC PROBLEMS OF ORDINATION

Problems common to essentially all ordination techniques have been widely discussed (Swan, 1970; Austin and Noy-Meir, 1971; Beals, 1973; Gauch, 1973; Dale, 1975; van Groenewoud, 1976; Austin, 1976a; Noy-Meir and Whittaker, 1977). The severity of the two major ones is a function of the β-diversity or heterogeneity of the data set, that is, how different the samples are from one another.

The first problem in sociological ordination is the nonlinear relationship of species to the environment and to one another. The second is the truncation of quantitative values of plant species along an environmental gradient at zero; if an environment is unfavorable to a species, an even more unfavorable environment does not decrease the species value below zero. Figure 1a illustrates both the nonlinear relationship of species to the environment (the bell-shaped curve) and their truncation at zero, and Fig. 1b represents the nonlinear relationship of two species. Sociological ordinations generally treat each species as an axis. When this is done, the environmental gradient is folded over along the axis of either species A or B, and furthermore, opposite segments of the environmental gradient will be



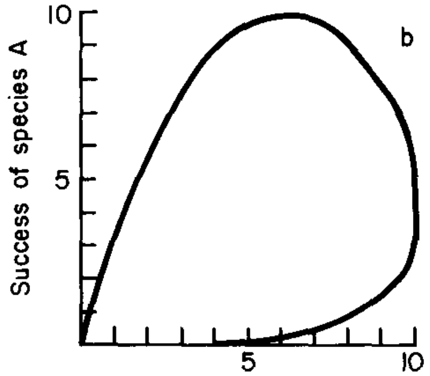


Fig. 1. Theoretical amplitudes of tolerance of two species which overlap along an environmental gradient (a), and the relationship of the species to one another (b).

combined and compressed to a single point (point zero) along the axis of either species A or B.

That environmental gradients appear in ordinations at all seems close to miraculous! It is only because that which is folded over or compressed by some species is extended by other species that information about environmental gradients is retrievable in sociological ordination. Some spatial models and some techniques may be more efficient at untangling these gradients, and some techniques that involve centering (as in covariance and correlation) are especially vulnerable to these problems.

Very similar problems are likely to occur in niche ordinations as well. Resource use may be nonlinear in the analogous sense that there will be an optimal resource type for a species, with less optimal resource types (because of less suitable characteristics, such as size or hardness) in a variety of directions from the characters of the optimal food item. It is also likely that many resource types will have zero values for many of the species under study. Everyone does not use the same resources. Unsuitability of a resource for a given species is truncated at zero. Therefore, although the emphasis here is on sociological ordination, most of what I say also applies to niche ordination.

All ordinations distort the original multivariate data set, and information is inevitably lost. There is a trade-off between loss of information and the simplification of data in order to detect pattern. Orloci (1974, 1975) has distinguished four types of distortion. Type A is that caused by forcing nonlinear data into a linear model. Type B is that of forcing a many-dimensioned configuration into a few dimensions. Type C is the distortion of using a "wrong" distance measure (Orloci believes there is a "true" distance). Type D is that of projecting a non-Euclidean space into a Euclidean space. There is an interplay of these types of distortion, which Orloci neglects—allowing one kind of distortion may reduce another, and the question is, which distortions are more tolerable, i.e., which obscure the ecological information the least?

Distortion in ordination has two kinds of consequences. The first is compressing and stretching distances in the ordination, compared with the original distance measures and relative to one another. This relates to Orloci's distortion types B and D; type B compresses many distances, while type D may compress or stretch distances. Orloci apparently thinks compressing distances a less objectionable distortion than stretching them; the latter may involve the occurrence of imaginary numbers or negative distances in the residuum after axis construction. But the loss of ecological information is similar in both cases.

The second consequence is the curvature of environmental axes, and this relates to Orloci's types A and C. This curvature has generally been considered a major distortion problem, which certain standardizations and distance measures can alleviate or aggravate. However, Allen (1981) suggests that curvature in ordination should not be viewed negatively but as a source of information. One source of curvature comes from bimodality of species—thus it reflects a real ecological phenomenon (the two ends of an environmental gradient may share some species) and can hardly be considered distortion. The other source is the result of nonlinearity and zero truncation of species, and thus curvature becomes a function of β-diversity. Both factors were involved in the rather circular moisture gradient in the ordination found in Beals and Cottam (1960). Allen provides numerous examples of bimodal species in Wisconsin vegetation. He also proposes that bimodality of absence is of similar ecological importance to bimodal presence. But "bimodal absence" of species (which increases in frequency as β-diversity increases or the environmental gradient lengthens) is not tied to specific environments as are modes of presence. And, whereas a species is likely to be bimodally present along only one environmental axis, absence will be polymodal along many axes. I think that inasmuch as gradient curvature reflects β-diversity, it is indeed distortion, however informative it may be. There are better ways to measure β-diversity.

IV. ALTERNATIVES TO BRAY-CURTIS ORDINATION

A. Principal Component Analysis

The poor performance of principal component analysis (PCA) is well documented on all but the most homogeneous sets of phytosociological data and ought to be relegated to the kind of continuous, linear, correlated, and nonzero data for which it was designed and for which it is excellent. Yet three major texts (Green, 1979; Orloci, 1975; Pielou, 1977) consider it the basic method of ordination of phytosociological data.

Empirically, its weakness for ordination has been demonstrated by Clymo (1980), Gauch and Whittaker (1972), Gauch *et al.* (1977), Jeglum *et al.* (1971), Kessel and Whittaker (1976), Lechowicz and Adams (1974), Mohler (1981), Risser and Rice (1971), Robertson (1978), Westman (1975), Whittaker and Gauch (1973), among others. Some (e.g., Risser and Rice, 1971) have found that PCA ordinations were totally uninterpretable, and that BC ordination made good ecological sense. Papers written in my class also attest to the inferiority and often unintelligibility of PCA for phytosociological or animal community data, and generally for resource-partitioning data (niche analysis) as well. Wolf (1975) has documented well the superiority of BC ordination over PCA for the last type of data.

Three alleged exceptions to these general results have been published. The first was Austin and Orloci (1966), who claimed PCA was 700 times more efficient than BC ordination, based on correlation of real distances between samples and ordination distances. Their evaluation was clearly unfair, however, because they used a non-Euclidean distance to construct the BC and tested its efficiency against a Euclidean distance. Furthermore, the patterns of species in the ordination, based on the limited data they present in their figures, are equally good if not better for the BC ordination. Mathematical efficiency is not the same as ecological informativeness (Austin, 1968; Austin and Greig-Smith, 1968; Beals, 1973).

The second exception is Walker (1975), who claimed PCA was better than BC ordination for his forest data. For PCA he used "zero-transformed" data, by which he eliminated zero values for absent species—a major problem in any ordination technique; for BC ordination he did not. Indeed, his comments imply that PCA on the same data as used for BC ordination was far less interpretable than the BC analysis. Thus, the superior performance of PCA was due strictly to the special data transformation.

The third exception is del Moral's evaluation (1980) of several different methodologies including PCA and BC ordination. Although I question some of his evaluation criteria, his one-axis-at-a-time analysis, and some circular reasoning, and although his presentation lacks the documentation to be really convincing, his data set may nevertheless exemplify an exception I would expect. If one has a large but fairly homogeneous set of samples, with a few samples which are very different from the rest ("oddballs"), BC ordination with its original choice of endpoints may not perform as well as PCA. del Moral's data do have low β-diversity, with a few oddball samples, and his worst performing BC ordination was based on the original endpoint selection. Using other criteria for endpoint selection, he apparently did make BC ordination as good as PCA in environmental interpretability.

We have also found that PCA may perform better than BC ordination when the data set consists of a small number of relatively homogeneous samples. Then, the severe limitation of available axes is a problem with BC ordination. Another group of data for which PCA seems equally good as BC ordination is the set of cyclic phytoplankton data through the seasons (Allen and Koonce, 1973; Bartell *et al.,* 1978). There is no evidence, however, that they are better (Bartell, 1973).

In addition to the empirical evidence against PCA, several theoretical arguments have been put forth (e.g., Beals, 1973). First, the use of correlation and covariance implies distances which emphasize ecologically less meaningful information, such as joint absences, and differences in rare species compared to differences in common species. In correlation more variable species are considered less reliable ecological indicators, whereas in reality variable species are more sensitive to ecological variation.

Second, rotating axes through a centroid that cannot exist in nature is a distortion of the environmental space and the sociological domain; no environment exists that would allow all species to coexist, except in very homogeneous data sets. Third, maximizing variance is questionable, first because clusters of samples are not hyperellipsoids but complex, multipronged configurations. When the two end portions of the first axis have residual variation based on entirely different species, for example, maximizing variance along a second axis effects a compromise that confuses the information within each end portion of the first axis.

But more importantly, in maximizing variance PCA necessarily tries to make all species correlate or covary linearly with each successive principal component as much as possible. (Specifically, it maximizes the sum of correlations squared for each species between the value of that species in a sample and the location of that sample along the axis.) Thus, PCA finds axes that do their best to make species peak at one end or the other, and it assumes that all species ought to do so, along some axis. Since along any extended vegetational or environmental gradient, many or even most species peak between the ends, PCA makes an unacceptable assumption about the ecological data and will distort whatever gradients occur in the data set.

A very short vegetational gradient with few species peaking between the ends may be considered primarily linear. Other data sets for which the assumptions of PCA are realistic include the sets of attributes based on community structure or dynamics, or based on morphological or physiological characters of species, all of which might be expected to vary unidirectionally along gradients. The reason why cyclic data are amenable to PCA, as suggested above, is probably that in a cyclic time sequence, samples will generally be near the periphery of the system (regardless of how many dimensions are involved) and hence all species will peak away from the centroid. Unlike an environmental gradient, which has two extremes, a cyclical time gradient is properly portrayed as a circle or other continuous line around a centroid. But for most sociological data and most niche data, PCA is really quite inappropriate and inevitably distorts the system. This problem is in addition to the distortion of using the Euclidean distance (discussed in Section V,C) which underlies PCA.

Arguments have been advanced (Boyd and Allen, 1981; Nichols, 1977) that PCA is still useful for data reduction, even if it does not give clear ecological patterns. However, PCA is based on the premise that covariance or linear correlation between attributes (species in our case) is a meaningful property. Indeed, a curvilinear correlation can be described by a linear coefficient, somewhat distorted but certainly not without meaning, but species–species and species–environment relationships are not even that. In Fig. 1a, an idealized relationship between two species along an environmental gradient is shown, a pattern widely documented and in its general outline irrefutable: species have upper and lower limits and optimal values along many environmental gradients. Figure 1b shows the interrelationships between the two species based on Fig. 1a. Random fluctuation is ignored, so the lines in Fig. 1a and 1b are the true parameters which a set of samples will approximate. It is entirely different from the parameters of the variance/covariance or correlation matrix assumed by PCA. Even for data reduction, I cannot see the value of PCA for this type of data; the basic information is necessarily severely distorted.

B. Other Eigenvector Methods

Other eigenvector methods have been tried. Factor analysis (Dagnelie, 1960, 1973), which maximizes communalities along each succeeding axis rather than variance, based on assumptions of underlying factors, requires considerably more work and is more arbitrary than PCA, and the results are similar, with the same inadequacies. It was used by Schnell *et al.* (1977) to look at tree distributions in Oklahoma; no species are modal there, so that linearity (or at least monotonicity) with environment (on a geographic scale) was more or less satisfied.

Canonical correlation is the simultaneous rotation of axes through species space and environmental space, so that one finds axes of best correlation between environmental gradients and compositional gradients. It has been tried often but found not particularly useful (Austin, 1968; Cassie and Michael, 1968; Cassie, 1969; Barkham and Norris, 1970; and Gauch and Wentworth, 1976), presumably because it is even more sensitive to nonlinear data than is PCA. Only Kercher and Goldstein (1977), using a relatively homogeneous data set (plots in one forest), thought it of much value. In any case, the method biases the phytosociological ordination toward environmental factors measured and thought important.

The ordination of Gower (1966), in which a symmetric matrix of similarity is calculated by any means of similarity or distance, implies a set of abstract variables which have Euclidean relationships (i.e., a matrix whose cross-products matrix is the matrix of similarity: find $\mathbf{B}$ such that $\mathbf{BB}' = \mathbf{S}$). Then PCA is performed on the new set of variables ($\mathbf{B}$). This allows the choice of a non-Euclidean metric for measuring distance or similarity. There may be distortions in the Euclidean representation, evidenced by negative eigenvalues. Orloci (1975) accepts this manifestation of distortion in a Gower ordination but strongly rejects exactly the same distortion as evidenced by imaginary residual distances in a BC ordination. Gower ordination is a possible improvement over direct PCA, but it still relies on an impossible centroid and on maximizing variance to extract information along each succeeding axis.

Several eigenvector methods have special functions. Although technically they are ordinations because they order samples (or species) along axes, they are not techniques to maximize overall ecological information (the usual role of ordination) but rather techniques to obtain specific information about the data set. Noncentered component analysis (Noy-Meir, 1973; Feoli, 1977) and varimax rotation [Kaiser, 1958; used in ecology by Ivimey-Cook and Proctor (1967), Noy-Meir (1971), Carleton (1980), and Wiegleb, (1980)] are designed to find natural groupings and thus they maximize information about gaps within the system. Discriminant analysis asks the question, what species best separate previously determined groups? It has been used in ecology by Norris and Barkham (1970), Grigal and Goldstein (1971), Goldstein and Grigal (1972), and Matthews (1979). It ignores information on similarities among sample groups and looks only at differentiating species. Furthermore, it treats a species uncorrelated with any other species as of equal importance as a group of many species strongly correlated with each other, and thus it obscures information on the number of species responding to specific environmental variation. It is therefore not an appropriate ordination for graphic display of overall vegetation patterns (see also Kessel and Whittaker, 1976).

C. Reciprocal Averaging

Reciprocal averaging (RA) or correspondence analysis [Benzecri, 1964; used first in ecology by Hatheway (1971) and Hill (1973)] may be viewed as an eigenvector ordination, but it can also be achieved for a first axis by a series of weighted-average operations. As an eigenvector method, it can be viewed as rotating axes simultaneously in species space and in sample space until the correspondence of each succeeding pair of axes is maximized. The end result is an axis in which the weighted averages of species produce the weightings (order) of samples and vice versa. It was foreshadowed (1) by Curtis and McIntosh (1951), who constructed a crude order of forest stands, ordered tree species along that gradient (though not by weighted averages), and then reordered the stands by weighted averages of species; and (2) by Bakuzis (Bakuzis and Hansen, 1959), who began with a very crude ordering of species along several environmental gradients, based on information in local floras, then ordered stands along each gradient by weighted averages of the species, ordered species by weighted averages of stands, and finally reordered stands.

Reciprocal averaging is a useful technique, and some believe it superior to BC (Whittaker and Gauch, 1978; Robertson, 1978), although its maximization function is rather esoteric. It has recently become very popular (Austin, 1976b; Bouxin, 1976; Noy-Meir and Whittaker, 1977; del Moral and Watson, 1978; Pemadasa and Mueller-Dombois, 1979; Sabo and Whittaker, 1979).

Reciprocal averaging has the satisfying property of perfect correspondence of a species ordination and sample ordination. The space implied by a system in which either samples or species can be represented as points is not defined, but perhaps it corresponds to the real environmental space in which samples and the optimal points or centroids of species occur. Another satisfying aspect is that it makes use of all information in the samples to construct each axis. An important advantage of RA is that it can handle longer gradients (more heterogeneous or with greater β-diversity) than can most ordination techniques, including BC as it is commonly used.

The disadvantages of RA, however, are substantial. It is good primarily for only one axis. Methodologies for subsequent axes are complex (Hill, 1973). The second axis tends to be an arch, and axes after the first are often hard to interpret (Hill, 1973; Bouxin, 1976; Gauch *et al.*, 1977; del Moral and Watson, 1978; Clymo, 1980; Marks and Harcombe, 1981; Prentice, 1980; Tyler, 1981; Persson, 1981).

As a specific example, Hall and Swaine (1976) used RA on a heterogeneous data set from Ghana forests. The first axis produced a useful gradient from the wettest forests to the outlying forest islands in savanna. But axes 2, 3, 4, and 5 all showed "polynomial dependence" on that first axis. The sixth axis was less dependent on the first and was used as the second axis of their ordination graph to get the best scatter of points. Environmental relations of the sixth axis were complex and inconsistent, and depended on where along the first axis the stands occurred. Some information probably did occur in axes 2–5, so that the sixth axis was not the best representation of the residual information after the first axis. In subsequent discussion Hall and Swaine devote their comments to trends along the first axis only. However, it is unlikely that all the important information was accounted for there.

Persson (1981) similarly discarded axes 2 and 3, which arched with the first, and used the fourth axis as the second ordination axis. He found a weak vegetational gradient associated with this fourth axis, but no clear environmental gradient.

If the data are responding primarily to a single gradient, the method may be a good one. If more than one gradient is involved, RA may combine uncorrelated environmental factors into one arbitrary gradient, and the potential independence of those factors may be irretrievable in later axes. These problems arise precisely because the information of all samples is forced to contribute to the first axis.

In one of our studies of forest vegetation, RA produced a first axis that was well correlated with moisture conditions from dry to mesic, especially for mature forests. Stands which represented rather late stages of secondary succession followed the same trend but were confined to the middle section of the first axis. When the first axis was examined for the distribution of mature forests only, those stands tended to be concentrated at either end, and the distances among dry-mesic stands were greatly stretched relative to those among either dry or mesic stands with comparable phytosociological distances. In other words, the presence of a second gradient of variation in the data rendered the gradient of primary variation distorted along the first ordination axis. It is clear from this why later axes form arches with the first. No later axes will correct for that earlier distortion. BC ordination gave a more reasonable distribution of mature forests along the first axis.

One reason that RA can construct a longer (more heterogeneous) first axis is that it takes into account species which occur in samples intermediate to the end samples but not in samples at the ends. BC ordination in its original form ignores those intermediate species in axis construction. However, RA cannot distinguish among (1) such intermediate species, (2) species found in some samples all across the gradient, (3) bimodal species found at both ends, and (4) species found in samples not environmentally intermediate but different from both ends. Species types 2, 3, and 4, unlike type 1, represent in various ways environmental variation not apparent between those end samples. Such information, lost in the first axis, may not reappear on later axes. Constructing axes from sociological distances (i.e., BC ordination) does distinguish types 1, 2, and 3 from one another, and type 4 from types 2 and 3, but not from type 1. Thus, certain types of information are lost in RA that are not lost in BC ordination.

Weighted-average techniques, including RA, put other limitations on the analysis. First, they automatically relativize data within each sample (in the sample ordination) and within each species (in the species ordination). There are no other options. Second, RA implies a system of Euclidean distances. Any weighted-average technique can be viewed as a system of Euclidean vector components. In locating samples, each species loading is proportional to the cosine of the angle of the axis in the dimension of that species, and the reverse is true in locating species in sample space. Thus, species (or samples) are necessarily considered Euclidean orthogonal axes, which is probably a contributing factor to the strong gradient curvature found in RA.

The comparative study of Gauch *et al.* (1977) indicates that BC ordination is superior to RA except when β-diversity is very high, whereas Whittaker and Gauch (1978) strongly favor RA over BC ordination, apparently without any further published evidence. More recently, Gauch and Scruggs (1979) admit that "some variants" of BC ordination perform better than RA.

Robertson (1978) applied RA and BC ordination as well as PCA, to forest vegetation. A look at his results is enlightening. Despite his claim that RA gave the best results, BC ordination and PCA actually gave a better spread of points in two dimensions. While the wet-to-mesic gradient is virtually as clear in BC ordination as in RA (but not clear in PCA), BC ordination does show considerable variation independent of that gradient, which Robertson seems to assume is noise. The question Robertson does not address is whether that second dimension is ecologically significant. His discussion implies that vegetational variation is based on a single environmental gradient, yet this is hard to reconcile with the scatter of points in BC.

One of his arguments that RA is better is that the stands were more equitably distributed along the first axis than with BC ordination. Actually this points up a further weakness of RA, i.e., that it may obscure real discontinuities in data sets. More generally, if along some portions of the gradient vegetation changes more rapidly than along other portions, RA is likely to obscure those differences. I have seen this trend in two of my students' analyses. This variation of response along a gradient is important information that I think should be evident in a sociological ordination.

In some instances in our work, RA has had the opposite effect, of exaggerating discontinuities. This seems to happen when few species occur across groups. Then, groups become tightly clumped and within-group information is lost. Occasionally distinct groups may be lumped tightly together. The effect of oddball samples on RA ordination is sometimes greater than on BC ordination. Under some conditions, the removal or addition of a sample from a set dramatically affected the ordination values of both species and samples, even when that sample is in the center of the ordination.

Mohler (1981) shows that RA does better when the ends of the gradient are oversampled compared to intermediate samples. However, to suggest this as a solution to RA distortions is to negate much of the value of a sociological ordination: the researcher must determine a priori the extreme conditions for all major environmental factors causing variation and search out such extreme field sites for extra sampling.

Hill (1979) and Hill and Gauch (1980) introduced a modification of RA called detrended correspondence analysis (DCA), which was used by Christensen and Peet (1981), Mohler (1981), and Sabo (1980). It eliminates any trend relationship between the first and second axes, including the arch, and hence presumably reflects new information along the second. It also reduces the compression of distances at each end of an axis. But with detrending, getting from the original spatial model to the final ordination involves intense manipulations that obscure the direct relation between that model and its simplified representation, the ordination. RA generates information (the arch) that is not wanted, so that information is obliterated by drastic measures. It is rather like taking a hammer to pound out an unwanted bulge. The question of why RA causes severe curvature if not given such special prophylaxis is avoided rather than answered: can information lost by compression onto the first axis ever be retrieved along later axes? The theoretical as well as practical ramifications of such treatment need to be evaluated. Furthermore, a real sociological curvature, due to bimodality of some species, will be eliminated by this method, and hence useful ecological information may not be detected.

In fact, one of the major selling points of correspondence analysis, the reciprocity of species and sample ordinations, is lost. In addition, Hill and Gauch (1980) admit that DCA still suffers, as does RA, from sensitivity to oddball samples and from poor estimation of discontinuous gaps in the data set. Unfortunately, they do not compare DCA with BC ordination. Mohler (1981) also shows that DCA, like RA, does better when extremes of a gradient are oversampled. Despite all this criticism of theoretical aspects, students have found that DCA can produce reasonable ordinations at times. Also, our experience indicates that a third or fourth DCA axis may contain more ecological information than a second detrended axis.

DCA has not been compared with BC ordination in the literature, but Gauch (1982) applied it to a subset of Bray and Curtis' original data and discovered that the second axis does not spread the stands out. He thus concludes that there is only "a single ecologically meaningful community gradient," a conclusion not supported by Bray and Curtis (1957). Curiously, Gauch's second DCA axis of the species ordination did pull out two tree species, *Carya ovata* and *Juglans cinerea,* both of which are ecologically rather peculiar in Wisconsin forests. Clearly, DCA and RA are very sensitive to oddball and rare species as well as to oddball samples, and they are therefore especially prone to obscuring the overall pattern.

Most published evaluations of RA and DCA have been based on artificial gradients or on real data with a single dominant gradient. The effects on these methods of oddball samples, of removal and addition of samples and species from or to the analysis, of highly skewed species distribution curves, of discontinuities, and of multidimensioned systems need to be studied more systematically. In any case, neither RA nor DCA is unequivocally the best ordination method.

D. Iterative Stress Minimization Techniques

Several techniques designed to minimize some measure of stress in an ordination are potentially effective. Stress is based on the distortion of compressing or stretching distances. It can be measured in a number of ways, but the subsequent strategy to find the minimum stress usually employs the method of steepest descent, an iterative computer technique using partial derivatives. Best known is Kruskal's (1964) nonmetric scaling (used in ecology by Anderson, 1971; Fasham, 1977; and Matthews, 1978), a method which tries to match the rankings of distances between sample pairs from the original data with the rankings of the respective distances in the ordination. Fasham (1977) found that Kruskal scaling gave better results than RA and PCA. A closely related technique is that proposed by Sibson (1972; used by Prentice, 1977, in ecology). Kelsey *et al.* (1977) proposed a technique based on the same principles, although the strategy was different.

A theoretical advantage here is that ranking linearizes the relation between environmental distance and sociological distances. Stretched and compressed distances are usually weighed equally. A disadvantage is that by not considering actual distance values, important information may be lost; there may be more than one equally good configuration based on rankings, and it is likely that discontinuities will be masked (Orloci, 1975). Anderson (1971), Wolf (1975), and some of my students have shown evidence for this distortion. On the other hand, Prentice (1977) found discontinuities preserved. Prentice (1980), using simulated data along one gradient, found that these methods could produce very contorted gradients, but, applied to simulated multiple gradients and to field data, they work very well.

Anderson (1971) proposed a metric equivalent to Kruskal scaling, minimizing stress measured quantitatively. His technique can be weighted to match more closely the shorter or the longer distances. I think his method has great potential and should be further tested. Another technique called "parametric mapping" or catenation (Shepard and Carroll, 1966; used by Noy-Meir, 1974) emphasizes closest adjacent distances in its stress minimization, and it has promise. The above iterative trial-and-error techniques all involve substantially more computer time than BC ordination, RA, and PCA. The method of steepest descent often risks finding a local minimum rather than the overall minimum, so that with these stress minimization techniques, one may not be certain the best configuration has been found.

Other iterative techniques include Gaussian curve fitting (Gauch *et al.*, 1974; see also Ihm and van Groenewoud, 1975), but these methods involve inordinate computational load, they assume that all areas of a gradient are equally represented in sample, and they assume symmetric, unimodal bell-shaped curves for all species. Polynomial curve fitting of a PCA ordination has been proposed (Phillips, 1978), but it cannot distinguish between real curvature of a gradient (due to bimodal species) and artificial curvature (due to nonlinearity). None of these last mentioned techniques has shown better results than simpler methods, and their improvement on even theoretical considerations is not firm.

V. THE BRAY-CURTIS METHOD

A. Definition and Strategy

Bray and Curtis (1957) rejected PCA (as used by Goodall, 1954) because it was inappropriate to phytosociological data, and they developed their own method. The essence of their method (BC) is (1) to calculate a distance matrix, (2) to select two reference points (either real or synthetic samples) for determining direction of each axis, and (3) to project all samples onto each such axis by their relationship to the two reference points. There have been many modifications proposed of the specific methodology originally used in the 1957 paper (Beals, 1960, 1965a,b, 1973; Gauch, 1973b; Gauch and Scruggs, 1979; Gauch and Whittaker, 1972; Maycock and Curtis, 1960; Orloci, 1966, 1974; van der Maarel, 1969; Monk, 1965; Swan and Dix, 1966;

Swan *et al.*, 1969, etc.). I consider all of these simply variants of BC ordination, as long as the above three essentials are followed. BC has been termed "polar ordination" (Goff and Cottam, 1967) or "Wisconsin comparative ordination" (Cottam *et al.*, 1973). Yet, it is polar only if the reference points are at opposite ends of the point cluster, which they need not be, and it does not seem to be more comparative than other methods.

There has been severe criticism of the method. Neither Pielou (1977) in her textbook on mathematical ecology nor Green (1979) in his textbook on statistical methods for environmental biologists deems BC ordination worthy of mention. Orloci (1975) devotes considerable effort to enumerating its weaknesses in his book, calling it "the least recommendable" method. Gauch (1982) is less critical and admits that "polar ordination remains of interest," although he strongly favors DCA. Some criticisms are unwarranted (Beals, 1973), others are easily corrected, and the remainder must be weighed against the weaknesses of other methods. Orloci (1975) did not take into account modifications published in 1965 (Beals, 1965a).

Many have said or implied that BC ordination is a crude approximation to PCA (e.g., Lambert and Dale, 1964; Greig-Smith, 1964) or is less rigorous or more informal than PCA (e.g., Goodall, 1970; Whittaker and Gauch, 1973). On the contrary, BC ordination is not necessarily an approximation of PCA in any sense, and while mathematically simpler, it can be as mathematically rigorous and as precise a method of ordination as any eigenvector method, and more so than the trial-and-error stress minimization techniques.

The remainder of this article discusses alternative procedures for BC ordination and responses to specific criticism, and proposes possible improvements.

B. Data Adjustments

The first decision to be made in the ordination of a data set is whether and how to adjust the data points. Should raw quantitative data be used, or some relativization or equalization? Or should presence/absence data be used? A few comparative studies have been done (Allen and Koonce, 1973; Austin and Greig-Smith, 1968; Austin and Noy-Meir, 1971; Noy-Meir *et al.*, 1975; Smartt *et al.*, 1974, 1976; Gauch and Scruggs, 1979), but much more work is needed. We have worked with a wide variety of data transformations without gaining much new insight. What is remarkable is that a wide range of standardizations of quantitative data often yields very similar results in the final ordination. This is because there is so much redundant information in a phytosociological data set; all species are responding to much the same environmental factors. However, some species may be better informers than others, and some measures of species success may be better informers than others.

Standardizations generally have one of two ecological functions. First, they may remove or reduce the effect of total amount of vegetation from the distance measure and subsequent ordination. Data are relativized by sample, so that $\Sigma\, x = 1$ or $\Sigma\, x^2 = 1$ for each sample, or the species values may be maximum-adjusted within the sample so that the most abundant species value equals 1 in every sample. Judged from our experience, these adjustments produce important differences in ordination results, compared with raw data, only when the sum of quantitative values for the lowest and highest stands differs at least twofold. The same environmental information reflected in quantity of vegetation is generally also reflected in species proportions. When they do give substantially different results, unrelativized data generally lump together on the first axis those vegetation-poor samples that are compositionally very different, whereas relativized data keep such samples separate. The second axis with unrelativized data often resembles the first axis with relativized data.

Second, standardizations may equalize the importance of all species to some degree. The assumption is that uncommon species may have as much to say about the ecology of the system as do common species. Logarithmic or square root transformations, use of presence/absence, relativization by species so that $\Sigma\, x = 1$ or $\Sigma\, x^2 = 1$ for each species, and maximum-adjustment so that the maximum value of each species among all stands equals 1, all tend to equalize species contributions to the ordination. They tend to emphasize species diversity in the subsequent distance measure as well. These adjustments, including presence/absence, may produce somewhat different ordination results compared to raw data or to within-sample adjustments.

Bray and Curtis (1957) introduced a double standardization that served both functions: first each species is maximum-adjusted to equalize species contributions, and then samples are relativized to reduce the effect of differing summed quantities. The result is an esoteric quantitative value, the implications of which have not been discussed in the literature, despite its frequent use. A high value indicates that a species is nearer its optimum in that sample relative to other species in that sample. The data matrix structure is highly modified and there is no monotonicity within either rows or columns between the adjusted matrix and the raw matrix. Nevertheless, comparative studies (Austin and Greig-Smith, 1968; Gauch and Scruggs, 1979) indicate that this double standardization gives more satisfactory results than raw data, sample relativization alone, or species maximum-adjustment alone.

In any case, if quantitative data are used, I recommend relativizing within samples for two reasons: it eliminates a first ordination axis reflecting quantity alone, which may lump unlike extreme samples together, and it qualifies the Sorensen distance measure as a metric. Presence/absence data, however, may be preferred for heterogeneous data sets (see van der Maarel, 1969; Gauch and Whittaker, 1972, etc.). They may also be relativized to reduce the effect of diversity on the first axis and to make the distance measure metric.

A third ecological function of data adjustments, which is much less often seen, is to correct the zero data points. The values are derived by looking in some way at the whole community structure contained in the data set. Since these adjustments are used primarily to extend distance measures, they are discussed later under that topic (Section V,D).

C. Distance Measures

All distance measures based on sociological data are subject to distortion compared with their corresponding environmental distances. If true environmental distances were known, it would be far better to do an environmental ordination. But, no ecologist has ever measured all the relevant environmental factors, including historical factors, and the importance to the community of those measured is seldom known precisely. Nor is the ecologist likely to understand the different sensitivities of the community in different parts of the environmental gradient. If these statements are true of physicochemical factors, they are even more true of the biotic factors that influence community pattern. Therefore we generally rely on the sociological data for an integrated measure of ecological distance between communities despite the distortion.

As environmental distance increases, sociological distance becomes less sensitive to environmental differences. The reasons are based on the two major problems already described, of nonlinear and zero-truncated relationships between organisms and their environment. To illustrate, assume a single environmental gradient with many species success curves along it, sampled at frequent intervals. If samples close together on the gradient are compared, they will share most species, and few species will reach their optimum between them; thus the difference between the two samples for most every species is essentially linear with the environmental difference. But if samples farther and farther apart along the gradient are compared, more and more species will reach their limits between them, and contribute no further information about environmental differences as more distant samples are compared. Furthermore, more and more species will also reach their optimal value between the samples, and the difference between the samples for those species will no longer reflect fully the environmental distance. (Two samples which have identical quantities of a species may be on opposite slopes of that species' curve.)

Thus, the nonlinear relationship of species and environment and zero truncation put an increasing limitation on the ability of sociological distance to reflect the environmental distance as the latter gets larger. The resultant curvilinear relationship between sociological and environmental distance was first described by Whittaker (1960, 1967). Beals (1973, Fig. 3) and Gauch (1973a, Fig. 1) have illustrated it for elevational gradients, and Swan (1970, Figs. 2, 4, and 5) did so for an artificial coenocline. It is seen in this article in Fig. 2 which is discussed in detail later in this Section.

In this regard, presence/absence data reduce the distortion substantially; the two-point distribution is symmetrically truncated, and the number of species shared by two samples reflects inversely but closely their distance along the environmental gradient. Only when two samples are far enough apart so that many species occur between them but not in them, or when the number of species in samples varies dramatically, will distortion become noticeable.

A major consideration in choosing a distance measure is its sensitivity to these distorting effects, and a second consideration is its sensitivity to sampling error.

Bray and Curtis (1957) originally applied a quantitative version of the Sorensen coefficient of similarity:

$$C = 2p_{jk}/(p_j + p_k)$$

where p_j and p_k are the sums of species values for samples j and k (for presence/absence data simply the number of species in each sample) and p_{jk} is the sum of the lesser species values for those species common to both samples (for presence/absence data simply the number of species in common). They used it as a percentage ($\times$ 100) and subtracted it from 100 to get a (percentage) distance. Curtis later felt that, because if one sampled a stand twice the coefficient would be less than 100, it would be preferable to subtract the coefficient of similarity from some value less than 100, namely the maximum value of the similarity which occurs in a matrix ($D = C_{\max} - C$). The maximum similarity value can be based on resampling the same community several times. This procedure makes the estimate of distance a conservative one. Maycock and Curtis (1960), Loucks (1962), and Beals and Cottam (1960) use this procedure, although they do not specifically state as such; instead, the information is contained in the theses upon which their papers were based. Beals (1960) first detailed the procedure in the literature, and it has been used by McIntosh and Hurley (1964), Austin and Orloci (1966), and Bannister (1968), and was advocated by Cottam *et al.* (1973). These derivations of distance, as Orloci (1975) points out, can increase non-Euclidean distortion in an ordination; an overly conservative distance is more likely to require stretching. Although this is true, the evidence is that such distortion is not likely to obscure ecological meaning. Bannister compared the use of $C_{max} - C$ and $1 - C$ and found very little difference. (He used 1 rather than 100 because his data were expressed as proportions not percentages.) Because sampling error is almost as likely to cause an underestimate as an overestimate of true distances, except when samples are very similar, $1 - C$ is generally a less biased estimation, as well as theoretically more desirable (see also van der Maarel, 1969). I see no advantage to $C_{max} - C$. Furthermore, $1 - C$ (Sorensen distance) can be calculated directly as

$$D_{hi} = \sum_j |x_{hj} - x_{ij}| / (\sum_j x_{hj} + \sum_j x_{ij})$$

where D_{hi} is the distance between samples h and i, and x_{hj} and x_{ij} are the values of species j in samples h and i, respectively. When $\Sigma x_{hj} = 1$ for all samples h, this simplifies to

$$D_{hi} = \tfrac{1}{2} \sum_j |x_{hj} - x_{ij}|$$

and it becomes simply a city-block metric.

Arguments for the use of city-block metric and allied measures in ordination, and against the use of a Euclidean metric, have been made previously (Beals, 1973). The advantage of the former is that all species contribute to the distance measure in proportion to their relative differences in the two samples. This necessarily weights an environmental factor according to the number of species responding to it, as well as to how dramatically they respond, and there is no exaggerated influence of big differences over small differences. The city-block metric conforms to the biological fact that the difference for most if not all species reflects differences in the entire set of environmental conditions between the two samples.

However, Orloci (1973, 1975) has criticized the Sorensen distance because of two "undesired properties": (1) scaling of the distance measure may vary from one sample pair to another, and (2) distance may fail the "triangle inequality condition." However, most ordinations employing this distance have used relativized data, and when data are relativized, both "undesired properties" disappear. Even if Σx_{hj} does not equal Σx_{ij}, and the distance is not metric, there is no empirical evidence that this is a serious burden to the method, as he claims.

Orloci (1974) advocates a Euclidean distance for BC ordination,

$$D_{hi} = (x_{hj}^2 - x_{ij}^2)^{1/2}$$

but in addition to the theoretical arguments given in Beals (1973), all the evidence in the literature shows that the Sorensen coefficient gives ecologically more interpretable results in multivariate analyses (Williams *et al.*, 1966; Bannister, 1968; Newsome and Dix, 1968; Gauch and Whittaker, 1972; Gauch and Scruggs, 1979).

Orloci (1967) saw some disadvantages of absolute Euclidean distance and proposed a "standardized" Euclidean distance. One can use his equation directly on any quantitative data set, or relativize the data set first by making $\Sigma x_{ij}^2 = 1$, and using the above equation. This standardization removes the limiting effect of total species quantities in samples on the maximum possible distance. With absolute distances, two samples with no species in common but with very little total vegetation will necessarily be closer together than two samples with no species in common but with high total quantities of vegetation. The former, totally unlike samples may even be closer than two high-quantity vegetation samples with a moderate amount in common.

This clear improvement, however, constrains all samples to the "surface" of a hypersphere (more correctly a quarter-hypersphere) because all samples have the same Euclidean distance from the origin (all species, $x = 0$). Therefore, the ecological space is no longer Euclidean but Riemannian, that is, a curved space. A Euclidean representation of this total space has meaningless additional dimensions that simply reflect curvature. Fortunately, ordination selects major axes of the system, and often this curvature is ignored. Orloci assumes that the chord (straight line) distance (using the Euclidean equation on standardized data) is the appropriate one, even though that distance goes through a space outside vegetation space. (It is like measuring the distance from New York to Melbourne through the middle of the earth.) If sample B lies directly between samples A and C in terms of vegetational differences, in this curved space, then D_{AC} is necessarily less than $D_{AB} + D_{BC}$. It would be more appropriate to use the arc distance,

$$D_{hi} = \cos^{-1} [-2(\text{chord distance})^2/\pi^2]$$

This is the true distance between points h and i in the curved vegetational space. A Euclidean representation then projects this curved space onto a plane, presenting the problems typically faced by cartographers. The results of this arc distance are almost identical to those of the chord distance, and since the latter is much more commonly used, that is the one I compare here.

I should also point out that the correlation coefficient r, used in some eigenvector techniques, implies a distance which is often considered Euclidean but which in fact constrains points to a semi-hypersphere and hence to a Riemannian spatial system. The chord (straight-line) distance is

$$D_{hi} = \cos^{-1} [2(1 - r_{hi})]^{1/2}$$

whereas the arc distance is

$$D_{hi} = \cos^{-1} r_{hi}$$

We have tried correlation distances in BC ordinations and found them to perform badly. Bray and Curtis (1957) originally discarded correlation on theoretical grounds. The theoretical weaknesses are described above in the discussion of PCA in Section IV,A.

Not one class report by my students has ever demonstrated any Euclidean or Riemannian distance to be better than the Sorensen coefficient, and often it is much worse. Empirically and theoretically, Euclidean-type distances do not reflect the ecological system well.

Using my vegetation samples from the Ethiopian Rift Valley, for which I tried to minimize all environmental variables but elevation (Beals, 1969b), Fig. 2 shows the relationship of four phytosociological distances to environmental distance measured by elevational difference. Relatively few real data sets are available for which we know the true environmental distances, but elevational gradients may come close.

The critical aspect of the graphs in Fig. 2 is the environmental distance not in absolute terms but in terms of the β-diversity: at what environmental

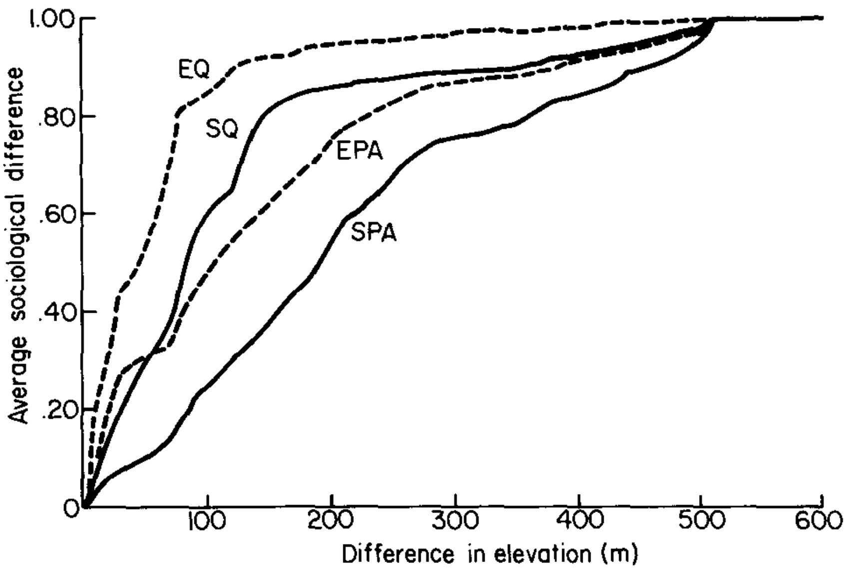


Fig. 2. Relation between elevational difference and four sociological distances: EQ, Euclidean distance, quantitative data standardized; EPA, Euclidean distance, presence/absence data; SQ, Sorensen distance, quantitative data relativized; SPA, Sorensen distance, presence/absence data.

distance do two samples have to be to have no species in common, and how straight is the line between that point and two samples with identical environments? The curvature will depend on the relative number of narrow-amplitude and broad-amplitude species in the system as well as on the measure used.

The result of the curvature and maximum truncation exemplified in Fig. 2 is that environmental gradients will be curved or bent in species space. That curvature is directly a function of the curvature (and angle of truncation) of sociological distance against environmental distance. It is also consequently a function of the length of the environmental gradient, i.e., the amount of β-diversity. All distance measures therefore will produce this distortion in the species-dimensioned space. On the other hand, ordination reduces dimensionality and may eliminate the curvature if the real vegetational gradients have greater variation than the induced variation (height of the arc of the curved gradient). The robustness of ordination is its ability often to ignore the curvature, and clearly the most robust technique will use the distance causing the least curvature, because then curvature is least likely to take precedence over real variation (ecological information) in constructing ordination axes. It is this ability of ordination to straighten out curved gradients in the total space that makes the use of such an environmentally distorted system at all tolerable.

The Euclidean distance is much more subject to distortion (curvature) than is the Sorensen distance (Fig. 2). With simulated data sets, the same relationship has been found every time. Euclidean distance is more sensitive to narrow-amplitude species than to broad-amplitude ones, since big differences in species between two samples are emphasized (by summing squared differences). Between two stands there will be more big differences for narrow-amplitude species than for broad-amplitude species, but this ratio will decrease as environmental distances get larger. Thus, Euclidean distances necessarily rise more sharply at smaller environmental distances and level off more dramatically as environmental distance increases.

We have applied several other distance measures to this Ethiopian data set as well as to artificial data: absolute Euclidean, absolute city-block, correlation, the Jaccard coefficient, Mahalanobis' D. None gave as near-linear results as the Sorensen distance. Absolute distances and correlation distance declined at larger environmental distances. The Jaccard measure, which is closely allied to the Sorensen distance and which has been recommended on the grounds that it is metric, gave more curvature than standardized Euclidean. Also, there was virtually no difference between relativized data and raw data for the standardized distances, even though in the case of the Sorensen distance one measure is metric while the other is not.

Although the advantages of a city-block distance measure are substantial, there are disadvantages. One is conceptual: we do not feel comfortable with a space in which the Pythagorean theorem does not hold, with graphs in which distances between points cannot be measured by a straight line between them, with circles which appear square and have a circumference equal to $4d$, etc. Second, city-block space is not amenable to direct matrix manipulation. One can convert it by a Gower ordination, but then all species values disappear into a framework of synthetic and virtually meaningless variables.

Third, axes cannot freely rotate in city-block space as they can in Euclidean space. On a Euclidean plane, a configuration of points retains its shape and its interrelationships regardless of the coordinates used to define their location. But this is not true on a city-block plane. A city-block system is illustrated in two dimensions in Fig. 3. The distance between points A and B, which is 10 blocks along one "street," is qualitatively different from the distance between points C and D, which are also 10 blocks apart, but 7 blocks in one direction and 3 in another. The shortest distance between points A and B can be achieved by only one straight line, but that between points C and D can be achieved by more than one "straight" line (defined as the shortest distance between two points). One can go 7 blocks east and then 3 blocks south, or 3 blocks south and then 7 blocks east, or alternate going east and south in a variety of ways, and still travel the shortest distance. In vegetation space, these blocks are infinitesimal, and so there are an infinite number of ways to get from sample C to sample D by the shortest distance.

To "rotate" axes so that the two points C and D are considered a straight-line coordinate (a single dimension) is to change the configuration and cause potentially serious distortion. Since it is not truly a rotation of axes, I prefer the term "deangulation" to describe the process of using the distance between two points in many dimensions as a coordinate of one dimension. Several points in Fig. 3 could have identical distances from both points C

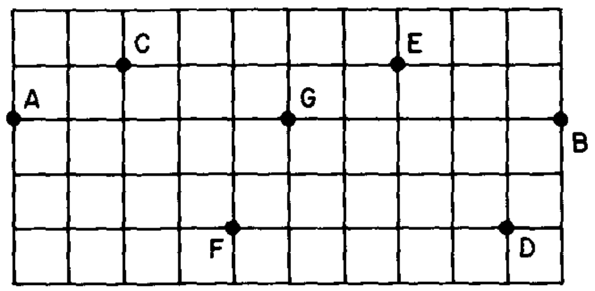


Fig. 3. An example of a city-block system in two dimensions. The distance between A and B is quantitatively the same as that between C and D but qualitatively different. Points E and F are the same distances from both C and D and yet are not the same point. G is the only possible point equidistant from A and B.

and D. For example, points E and F are both 5 blocks from C and 5 blocks from D. If we consider the distance between C and D as a coordinate 10 blocks long, E and F will both fall at the midpoint of the CD coordinate, even though they are themselves 6 blocks apart. This collapse of a universe of straight lines could be disastrous, were it not for the fact that not all of the theoretical city-block space is occupied by real points (not all species combinations are possible). Most species are correlated (in the broad sense) and show some kind of trend (Fig. 1) with most other species. Therefore, if point C is low in species X and low in species Y, it is unlikely that another sample will be high in both species. The real data set has many more dimensions than the two in Fig. 3, and a sample containing species common to two far apart samples will generally be intermediate in species values between those two far apart samples (close to an imaginary diagonal from C to D in Fig. 3). It is clear from most studies that the distortion thus induced is more than compensated for by the reduced distortion in the distance measure itself.

However, the use of city-block space sooner or later requires translation into a Euclidean space for graphic display of an ordination, and for correlation and regression analysis. This inevitably leads to further distortion, but this distortion actually counterbalances the distortion of deangulation and reduces the total distortion, even though imaginary components to distance may appear in the residual distance matrix. The distance between E and F, lost by deangulation of the CD axis, may be restored in the projection of city-block space into Euclidean.

Finally, although the city-block metric reduces the distortion of environmental distance, compared to other measures, it does not eliminate that distortion. Methods of reducing it further are discussed in Section V,D.

Figure 2 also shows that presence/absence (p/a) data give less distorted results than quantitative data for the particular vegetation system used. The pattern holds for our artificial data sets as well. van der Maarel (1969) emphasizes the better behavior of p/a data for complex vegetation systems.

However, the variance of the Sorensen distance at any given elevational difference is 30–80% greater for p/a data than for corresponding quantitative data. There appears to be more random variation in the former. Thus if the environmental range is within the nearly straight-line portion of distance measures for both p/a and quantitative data, the latter give more reliable results. Generally one does not have this information directly, but if most of the distance values are less than 0.80, the system is probably within that environmental range.

If p/a data are not relativized, the Sorensen distance is not a true metric, although it gives the least distortion in Fig. 2. One could relativize the data, in which case the value of each species in a sample is the reciprocal of the total number of species in that sample. The results are nearly identical with those of unrelativized p/a data for the data sets observed. Or, one could use the absolute city-block distance, i.e., the number of species in one or the other samples but not in both. This is equivalent to the information statistic of Williams *et al.* (1966). However, this measure can decline with larger environmental distances, and van der Maarel (1969) reports that it is less efficient than the Sorensen distance in systems with high β-diversity. Incidentally, absolute city-block distance is simply the square of the absolute Euclidean distance when p/a data are used.

Finally, there is another complication to the exclusive use of p/a data. The elevational series in Fig. 2 is dominated by a single (though multifactored) environmental gradient. Our simulated coenoclines were also one-dimensional gradients. When plant or animal communities vary along more than one gradient, as they generally do, some aspects of environmental variation may be better detected by p/a data and some by quantitative data. Several studies (e.g., Allen, 1971; Allen and Koonce, 1973; Wolf, 1975) have found that both types of data from the same set of sample points produce substantially different but equally informative ordinations.

The question of which kind of data to use to calculate the distance measure remains open. For very heterogeneous or very homogeneous data sets, the answer may be clear (p/a for the former, quantitative for the latter), but most data sets might benefit by having both types of distances analyzed. In sum, the most effective distance measure is certain: Sorensen distance gives a more nearly linear correlation with environmental distance than any other measure, whether quantitative or p/a data are used, or whether raw or relativized data. The problem of nonlinearity is reduced but not eliminated, and the problem of maximum truncation remains.

D. Straightening and Extending Distances

The curvilinear nature of sociological distance and its truncation at a maximum distance are major limitations to ordination, and the latter especially is a reason why current techniques are inadequate for very heterogeneous data sets.

There are some distance measures which are not truncated at a maximum value when two samples share no species (e.g., absolute city-block and Euclidean, correlation distances). The actual value of the distance when two samples have nothing in common varies on the basis of the total quantity of species in each stand, which generally has no bearing on ecological distances. (Two totally dissimilar samples, both of which have lots of vegetation, are not necessarily more different environmentally than two totally dissimilar samples with sparse vegetation.) Furthermore, the mean distance may not be monotonic with environmental distance from a given sample, if farther samples have fewer species than do closer samples. Therefore, these distances are even less adequate than the measures which are truncated at a maximum. These latter measures (such as Sorensen distance) are then preferred.

We have tried several strategies to overcome the two basic limitations of these distances. One is to adjust the data so that the zero problem is eliminated. Such an approach, however, does not solve the nonlinear problem, and in fact may increase it.

Two such adjustments have been suggested in the literature. The first is a double standardization of rows and columns of the data matrix (Austin and Greig-Smith, 1968); this I call "contingency deviate." It expresses a species quantity as a deviation from the value expected if all species were independent of the samples and of each other. Absent species will have varying negative values, as will those with low presence. The degree of absence (measured by the negative value) increases for species which are abundant elsewhere in the study and/or which are in samples with high species totals. The former emphasis might have some ecological value but the latter does not. In fact, a species could easily have a lower value in a sample in which it is present than it does in another sample in which it is absent.

A more ecologically realistic attempt to quantify absence was made by Swan (1970). An association index between all species pairs was calculated. For a species absent in a sample, Swan calculated its average association with all species present in that sample. Thus, the species in the samples indicated the likelihood that the absent species could have occurred in the sample. These values ranged from 0–100, and for any species present in a sample, its actual value was added to 100 to give it an adjusted value above the range of absences. This adjustment has been used by Jesberger and Sheard (1973) and Walker (1974, 1975) among others. Its major drawback is that the scales of absence and presence are not comparable. One is measured in terms of joint occurrences with other species, and the other is measured by some observed quantity about the species itself.

We have tried a similar measure, the "sociological favorability index," generated entirely from presence/absence data:

$$b_{ij} = (1/S) \sum_{k}^{S} N_{jk}/N_k$$

where S is the number of species in sample i, N_{jk} is the number of samples with both species j and k, and N_k is the number of samples with species k. The term b_{ij} is the average probability of species j being present in sample i, judged from all species k in sample i. It reflects the favorability of the environment of sample *i* (biotic and presumably physicochemical) for species *j*. Species *j* is considered as one of species *k* if it occurs in sample *i*, and thus contributes either 1 or 0 to the averaged probability, depending on whether it is actually present or absent.

This has been our most effective solution to the zero-truncation problem. It extends the distance measure considerably: maximum value occurs when none of the species in sample A ever occur with any of the species in sample B. It reduces curvature, compared with the quantitive Sorensen distance, although it is more curved than when p/a data are used (but it nevertheless extends the distance).

The distance measure omits information on the quantitative value of each species, but it uses alternate information on that species' relationships to all other species instead. We have compared the sociological favorability index with actual density data for over 50 species from three data sets, and found that the correlation *r* (using only samples in which the species is present) ranges from 0.82 to 0.97, with a mean of 0.92. Thus the index, though derived from p/a data only, estimates the species quantities rather well (not in absolute terms but within each species' range of values).

A second, obvious strategy for the curvilinear problem is to straighten the curves in Fig. 2 by some mathematical transformation of the sociological distance. With the data from Fig. 2 and our artificial data, we were indeed able to approximate a straight-line relationship between Sorensen and environmental distance by various trigonometric and polynomial transformations, but each data set required a different transformation, depending on the proportion of wide- and narrow-amplitude species, etc. To do this correctly, we would need to know more about the data set than is normally available. Gauch (1973a) has given a theoretical transformation, which assumes that the gradient includes mostly species with symmetric Gaussian distributions of similar width, a highly unlikely real world situation (see also Gauch, 1973b). The distances after transformation are of course no longer metric, but a more serious problem is that the variance of the higher values is greatly exaggerated. Also, transformations do not extend the distance.

A trigonometric transformation of Sorensen distances has been made by Loucks (1962), Austin and Orloci (1966), and Gauch and Scruggs (1979), using the arc sine of the square root of the original distance. The adjustment is recommended in statistics to make the mean and variance of proportional data independent. That rationale is irrelevant, however, for these distance measures, because they are not proportions of a sum grand total (that is, all distances do not add up to one) and hence the mean and variance are already independent (even assuming that that is a desirable quality of the distance matrix). But all three studies do suggest slight improvement of ordination by using the arc sine transformation. Gauch and Scruggs (1979) claim that this transformation approximates an error function based on the relationship between Sorensen distances and sample separation along a gradient. This is not true; if one transforms the Sorensen curves in Fig. 2 by arc sine, the distortion (curvature) will actually be increased for values under 0.50, although at higher values the curve is indeed straightened somewhat. I suspect the noted improvement is more coincidental than theory based. Perhaps the straightening of the upper part of the curve improves performance of the distance measure more than the increased curvature in the lower part worsens performance.

A third strategy that straightens the line and extends the distance measure is to incorporate in the measure some aspects of the vegetation or the animals that vary on a broader scale than do species quantities, possibly higher taxa such as genera or families, or structural aspects. I have done this with a lichen ordination (Beals, 1965a), in which I incorporated growth form as well as species. For woody vegetation, size classes of trees and shrubs and/or leaf shapes might be incorporated, and thus structural similarity is accounted for in the ordination as well as species similarity. Structural attributes are more likely to be linear with environmental variation than are taxa. Knight (1965) and Knight and Loucks (1969) have used structural attributes alone with some success, and van der Maarel (1972) used higher taxa (genus, family, order) exclusively to ordinate. This is an extremely powerful extender. Its primary weakness is that the proportion of distance alotted to structure or higher taxa vs species is highly arbitrary. The number and type of structural categories also will be arbitrary.

A fourth strategy which we have tried is a kind of second-order distance measure. First, a distance matrix is calculated, and those values are then used as attributes to calculate a second distance or correlation matrix. These distances then reflect the dissimilarity of two samples in their sociological relationships to all other samples. Originally we used a correlation matrix of the distances, but the results were quite unsatisfactory because of curvature, nonmonotonicity, etc. An absolute city-block distance measure works well to get the second distance matrix, and seems fairly good for straightening the distances out and for extending it slightly, but if there are many maximum distance values in the first matrix, an absolute distance measure may decrease at very large environmental distances.

The fifth strategy is to use a stepping-stone distance when two samples j and k are totally dissimilar. Find the sample i that is most directly between the two, and sum the two distances D_{ij} and D_{ik}, to get an extended distance between. That is, when $D_{jk} = D_{\max}$, calculate $D_{ij} + D_{ik}$ for all samples i ($i = j, k$) and find the lowest value; this is the extended distance $\hat{D}_{jk}$. When two samples are far apart, it is likely that there will be a sample almost directly between them. There will be some tendency to overestimate the true environmental distance, and hence the possibility of stretching other distances in ordination construction later.

If, in finding the lowest $(D_{ij} + D_{ik})$, either or both of those terms is still $D_{\max}$, then one must find two intervening samples, i and h, and the lowest $(D_{ij} + D_{hj} + D_{hk})$ becomes $\hat{D}_{jk}$. Theoretically, three or more intervening samples should be used if one of the terms in the above is still $D_{\max}$, but finding $\hat{D}_{jk}$ becomes more complicated. Unless the gradient being considered has three or more complete turnovers in species composition, it is unlikely that more than two intervening samples will be needed. This method is similar in principle to the step-across method of Williamson (1978) which dealt only with presence/absence data in preparation for Gower ordination.

I have used both second-order and stepping-stone distances on a data set from an unpublished study of vegetation in the Ethiopian Rift Valley, where the vegetation ranged from desert scrub to dense montane woodland. The use of stepping-stone distances was far superior and gave interpretable axes in two dimensions, whereas second-order distances gave a somewhat interpretable first axis and a confused second. Without such extensions, the ordination gave contorted environmental gradients. Originally I applied stepping-stone distances only when $D = 1$ (i.e., maximum), but it could be used when $D \geq 0.80$, to straighten out the curvature of the line in Fig. 2.

In summary, the best solutions to extending sociological distance measures to cover a wider environmental range seem to number three: (1) the use of sociological favorability index to represent the species in the data matrix, (2) incorporation of structural or higher taxonomic attributes in the data set, and (3) calculation of a stepping-stone distance when two samples are very dissimilar. The last two also tend to make the relationship of sociological and environmental distances more linear. The first two maintain whatever metric quality the distance measure originally had, while the last-one removes any metric pretense. On the other hand, stepping-stone distances can extend the distance further than the first two and straighten the relationship with environmental distance more effectively.

These extension methods are helpful primarily when there are more than just a few (for example, 5 or 10%) sample pairs with maximum distance. Otherwise, they hardly seem necessary.

E. Choice of Reference Points

Once the distance matrix is calculated, the next question is the choice of reference points for each axis. The choice is critical to the effectiveness of a BC ordination. If residual distances (discussed later in Section V,F,2) are used to construct each succeeding axis, as we strongly recommend, the criterion for choosing reference points is the same for each axis. Reference points may be real samples or synthetic points based on the average of several samples or some other hypothetical combinations of species (derived of course in an ecologically meaningful way). Several possible criteria are given below, with their advantages and disadvantages.

Whittaker and Gauch (1978) criticize the BC technique as "limited by their use of only a few samples to define axes." Thus, they feel eigenvector techniques, in particular RA, are superior because more information is used to define axes. The use of more information, however, is not necessarily better, as is clear in the case of PCA, which contorts both environmental and vegetational gradients, and in the case of second and higher axes of RA. Two samples may define a BC axis, but, in some way, the information of the entire data matrix is used to determine which samples will define the BC axes.

If one wishes to maximize information that can be achieved only be defining axes by all the sociological information (variance, heterogeneity, correspondence of species and sample axes, etc.), then the original BC technique is inadequate. However, certain kinds of information (such as range) are maximized only by using two samples as reference points.

The use of two samples risks loss of important information, but if chosen well may actually given clarity to the vegetational gradient, whereas other techniques may use "too much" information and give a more confused gradient. In any case, reference points are not limited to single samples but may be centroids of groups of samples.

1. Maximum Range

This is the original criterion of Bray and Curtis (1957). It is not, as many have suggested, some kind of approximation or subjective choice, but a perfectly legitimate, precise, and ecologically valid criterion to maximize the range of the samples along each succeeding gradient. The distance matrix is examined, and the two samples farthest apart are determined. If more than one pair have the same maximum value, then the first reference point is the sample with the highest sum of distances (that is, the one most different from all others). The second reference point is the sample most different from the first, or if there are more than one, the most different sample with the highest sum of distances.

Although this criterion is legitimate, in practice it often tends to concentrate samples in the center of the axis or at one end, isolated from both or at least one of the reference points. Maximizing range will separate out any oddball samples, and the ordination is more affected by such samples than is PCA (although oddball samples will shift the component vectors substantially). This has been a major reason for criticizing BC ordination (Austin and Orloci, 1966; Whittaker, 1967). The axes, however, do show ecologically useful information, i.e., that extreme stands are isolated from the rest in their species composition, but the information concerns a relatively small proportion of the samples.

To reduce the effect of oddball samples, they can be eliminated from consideration as reference points. We have found that a satisfactory approach is to eliminate all samples which have no distance less than or equal to the mean of all distances in the matrix. This provides not only the maximum range (excluding oddball samples) but also guarantees a good spread of points along the axis. There is, however, an arbitrariness in determining how "odd" samples have to be in order to be eliminated.

2. *Correlation*

van der Maarel (1969) proposed choosing as reference points the two samples which have the greatest negative correlation between their distances. Such reference points in theory should be at the ends of the long axis, and we have found that this criterion works well when there is a single major axis of environmental variation. But when there are two or more major axes or several oddball samples in the data set, the criterion becomes confused and may not select the samples at the end of the longest axis. Along succeeding axes, correlation seems even less satisfactory.

3. *Variance–Regression*

This criterion is similar to correlation criterion but overcomes to some extent the latter's sensitivity to oddball samples and secondary axes of variation. Correlation measures perfectness of relationship (how close points are to a straight line), while regression measures magnitude of relationship (slope of line). It has proved to be for us the most generally satisfactory criterion using real samples as reference points. It has been widely used at Wisconsin in theses (e.g., Bartell, 1973; Colburn, 1975; Denslow, 1977; DeJong, 1976; Howe, 1977; Kantak, 1977; Kline, 1976; Rusterholz, 1979; Shepherd, 1975; Thomson, 1975; Waide, 1973; Wolf, 1975; Wood, 1979; etc.) and published papers (Emlen, 1977; Lechowicz and Adams, 1974; Stephenson, 1974; Will-Wolf, 1980), although it has not been formally presented in the literature. It appears in a mimeographed paper by W. Post, E. Beals, and T. Allen, for use with a computer package at the University of Wisconsin. Work by my students has almost invariably found the variance–regression criterion to give more interpretable ecological results than other real-point criteria.

This method is formally presented here. The variance of the distances for each sample is calculated (excluding zero distances, i.e., distance with itself and, on succeeding axes, residual distances between reference pairs of preceding axes). The first reference point is the sample with the highest variance. This will be the sample at one end of the longest axis. Then that column of distances (for the first reference point) is compared with all other columns in the distance matrix, and the regression calculated of each other sample (column) against the first reference point. The sample that has the lowest regression value (i.e., the greatest negative value) is chosen as the second reference point. It will be at the other end of the longest axis. Regression is less affected than correlation by the scatter in points caused by oddball samples and secondary axes of variation. Basically, this method finds the longest linear axis, removing the first reference point from consideration. Oddball samples are automatically excluded unless they happen to be along the axis of major variation of the rest of the samples (in which case they are not a problem). The advantage of this method over the maximum range method with oddballs excluded is that there is no arbitrariness in what an oddball sample is. The question is not whether a sample is an oddball or not but whether it falls at the edge of the axis of major variation defined by the other samples.

As a result, this method tends to analyze complex-shaped clusters of data points (from horseshoes, or L and X shapes to those with multidimensional prongs), one linear axis at a time, and does not become muddled by combining unrelated axes, as maximizing variance invariably does, and as maximizing range may to some extent.

4. Extreme Environment

Whittaker and colleagues (Gauch and Whittaker, 1972; Cottam *et al.*, 1973; and Whittaker and Gauch, 1978) have suggested that, because maximum range so often reflected oddball samples, reference points could be chosen on some criterion external to their sociological content, namely environmental dissimilarity. In other words, the two samples that are environmentally most dissimilar along major axes of environmental variation are chosen. While I admit that better results are often obtained this way than with the original BC criterion, the major value of sociological ordination compared to environmental ordination is lost, that is, the freedom from assumptions about what the important environmental factors are. Maycock and Curtis (1960) used this criterion along the first axis of their ordination, using synthetic samples rather than real ones. They combined five samples at each end of the moisture gradient for their first two reference points. They were confident that moisture was the cause of the major axis of vegetational variation.

Even if confident about the environmental variation underlying the major vegetational gradient, researchers are much less likely to know the environment underlying secondary gradients. Maycock and Curtis (1960) reverted to sociological criteria for their second and third ordination axes.

One result of their first-axis procedure was a straightening of the moisture gradient more than it should have been based on vegetation alone, since a number of species in their data set were bimodal and occurred in both wet and dry sites but not in mesic sites. Comparing the ordinations of Beals and Cottam (1960) and of Loucks (1962) on data ecologically similar to those of Maycock and Curtis reveals that in both cases the moisture gradient is distinctly curved. In the first paper, the first and major axis went from extreme dry forests to mesic forests (not to wet) because they were the most different forests. The major axis of remaining variation ranged from dry-mesic to wet forests. Both of these latter forests shared some species with the dry as well as the mesic stands, but they did not share many of the same species with each other. Putting the two axes together resulted in a moisture gradient curved in a three-quarter circle, which was a far better representation of the vegetation pattern than a straight-line moisture gradient would have been. Of course, on the first axis alone, the pattern would have been obscure.

5. *Minimum Residual Distance*

Using a computer, all pairs of samples can be tried as reference points and that pair selected which gives some minimum stress to the ordination. Stress is usually defined in terms of the residual distance, the distance not accounted for by the ordination axes. For example, the pair of samples can be determined which maximizes variance along the axis, which in fact minimizes sums of distance-squared unaccounted for. This approximates PCA but is not constrained to a Euclidean metric. It is only an approximate method, and if used as a measure of stress, then PCA or a Gower ordination would be better.

But should stress be measured as sums of squares? Should a residual distance twice as great as another contribute four times as much to the stress measure, or only twice as much? If the latter is assumed, then the criterion of sums of residual distances (unsquared) should be used. For a Euclidean distance, such residual distances are calculated by the equation,

$$\mathrm{RD} = (D^2 - \mathrm{OD}^2)^{1/2}$$

where RD is residual distance, D is the original matrix distance, and OD is the distance in the ordination.

For a city-block distance, the residual distance for this purpose must be calculated as

$$\mathrm{RD} = |\, D - \mathrm{OD} \,|$$

(If not done this way, peculiarities of city-block geometry play havoc with axes after the first.)

In either case, this is an exact solution, not an approximation. That is, the axis through the cluster of data points which minimizes the stress (sums of residual distance unsquared) necessarily passes through two real samples. After spending considerable effort working out the algorithm for freely rotating axes in multidimensional space to find that minimum stress, I found that it had to go through two samples. It is much simpler, then, to calculate the stress for all pairs of samples and choose the pair with the lowest stress.

6. Perpendicular Axes

Orloci (1966) proposed an ordination (OPA) which is of the BC type but maintains exact relationships of all reference points to one another. Along the first axis, this allows two reference points as for any other BC ordination, but along each succeeding axis only one reference point can be added: for N axes there are $N + 1$ reference points. Orloci viewed the entire first axis as a reference for the second axis, the plane of the first two axes as a reference for the third axis, the volume of the first three as a reference for the fourth axis, etc. In practice, it is simpler, using residual distances described later, to consider the first reference point of the first axis as the first reference point for all later axes. Orloci used the criterion for choosing reference points as the most distant sample from the line, plane, volume, etc., defined by previous axes. This does not maximize range along each succeeding axis, although axes will be progressively shorter.

Orloci was trying to correct two alleged deficiencies in the BC ordination: nonperpendicular axes and the nonintersection of axes ordination with each other. These complaints were also made by Lambert and Dale (1964) and were repeated by Orloci (1974, 1975). I addressed those complaints previously (1973), but will do so again. Technically it is true that if two independent points are used for each axis, the line drawn between one set of two points will not necessarily intersect with the line drawn between another set. However, because any vector parallel to each of those lines will give the same ordination results, one can (and does) project those axes onto one another and the deficiency is spurious. The nonperpendicular axes problem is discussed later in this article.

In 1963 (unpublished) I developed the same technique as Orloci's perpendicular axes, and found it substantially less efficient than the original BC method. I calculated efficiency as correlation squared between ordination and original distances. Two OPA axes were only 72% as efficient as two BC axes; three OPA axes 60%. Looked at another way, two BC axes were 95% as efficient as three OPA axes, and three BC axes were 92% as efficient as five OPA axes. The reason for this reduced efficiency is that less information goes into constructing axes after the first; thus the information in four reference points is expressed in two BC axes but three OPA axes, etc.

Also, there are many fewer potential vectors to choose from after the first axis (Beals, 1973). For example, if there are 100 samples, then for the second axis regular BC ordination has 4753 vectors to choose from, whereas OPA has only 98; for the third axis BC ordination still has 4560 and OPA has 97. OPA is thus severely limited in the potential direction of ordination axes.

This method may have a value, however, in converting a non-Euclidean system into a Euclidean one. If one carries axis construction to $N-1$ dimensions by this perpendicular axes method, one has placed all samples within a Euclidean framework. There will be, in the residual distances, some imaginary numbers, but they will be small if each axis is the longest possible, given the remaining distances. If a PCA is performed on this $N-1$ OPA ordination, the results are similar to a Gower ordination. If all imaginary distances are set to zero, then the new distances calculated from the OPA can be used in any other BC ordination with Euclidean impunity. Those distances will be only slightly different from the original distance measures, and only different in cases in which there would otherwise have been a chance of nontriangularity to develop in Euclidean axis construction. I have tried this only with a data set of 20 samples, but the result was virtually identical to a BC ordination on the original distance matrix.

Swan *et al.* (1969) used a modified OPA, trying all possible reference points for each axis and choosing the pair of reference points for the first axis which gave maximum variance, and for remaining axes the individual reference point which gave maximum variance. This approximates PCA or Gower ordination, and has little to recommend it over those.

7. *Centroid and Reference Point*

This is Orloci's (1966) "position vectors" technique. The centroid (average of all samples) is one of the reference points for all axes. The other reference point is chosen to maximize variance (all samples must be tried to find it). This method was intended to approximate PCA, and is identical to the previous method except that the first reference point is the centroid and not a real sample, and the choice of additional reference points is specified differently. Thus, it has the same deficiencies as the previous method—few vectors to choose from and less information used per axis—but the deficiencies hold for the first axis as well as later ones. For example, with 100 samples, the first axis of regular BC ordination or of OPA has 4950 vectors to choose from, while position vectors has only 100. Furthermore the centroid will generally be well outside the domain of real points, and so the axis is forced to go through a nonexistent sociological space. Nevertheless, Orloci (1975) still prefers this method over regular BC.

8. *Functional Mean*

An alternative to the centroid technique would be to find the center of the cluster within real sociological space, rather than the centroid which is outside that space. This is the functional mean of Ramensky (1930). The sample most like all others (lowest sum of distances) will approximate this functional mean, that is, it will occur in an environment most intermediate or central to all other samples. We have tried using this criterion, but the results have been less useful than maximum range (excluding oddballs) or variance–regression. However, Sanford (1974) in his study of epiphytic orchids, used the most "typical" (functional mean) and least "typical" (most different) stands for reference points, and he suggests the results are more informative than had he used two stands with greatest interstand differences. It is not clear if he actually tried the latter for this data set.

9. *Synthetic Reference Points*

The problem of oddball samples was recognized from the beginning of BC ordination. To avoid it, Bray and Curtis (1957) actually used three pairs of reference samples for their first axis, two for their second, and only one pair for their third. (However, rather than average the three samples at each end for reference points of the first axis, they calculated three axes and averaged those). Maycock and Curtis (1960) combined five samples at each end of the first axis (based on extremes of a moisture gradient), though their succeeding axes were defined by single reference samples. The number and choice of samples averaged to produce a synthetic reference point are arbitrary, but these modifications can be an improvement over individual sample reference points.

The most powerful technique using synthesized reference points is to use all the samples in the production of those reference points. The samples are grouped in some ecologically meaningful way into two, four, six, or more groups, whose centroids then are points for one, two three, or more axes, respectively, of a BC ordination. This satisfies criticism (Whittaker and Gauch, 1978) that BC ordination uses only a few samples to define the axes. All information is considered in the entire ordination, but each axis selects a restricted amount of the total information, unlike RA. There are other advantages as well. These groups are less likely to have no species in common, and thus allow ordination of greater β-diversity.

The groupings may be achieved using any of the myriad clustering techniques. I have tried two (on several sets of data) that seem the best both theoretically and empirically (based on class data): sums of squares clustering (Orloci, 1967), an agglomerative polythetic technique which uses quantitative data, and association analysis (AA) (Williams and Lambert, 1959), a divisive but monothetic technique, which uses presence/absence data. The value of the Orloci procedure is that it finds groups which minimize within-group variance and maximize between-group variance. However, it utilizes the Euclidean distances among samples. The value of AA lies in its use of interspecific heterogeneity to group samples. Such heterogeneity is an obvious and meaningful property of the species' ecology. In the discussion below a BC ordination using Orloci's method is referred to as SS-BC and that using association analysis (Rusterholz, 1973) as AA-BC.

Having criticized the use of the centroid of a data set for ordination, I am now advocating the use of local centroids. However, these local centroids consist of relatively homogeneous samples, whose species generally do occur together. The environment expressed by the local centroids is somewhat specified. Thus they are at least on the edge of, if not within, the real sociological and environmental spaces defined by the data set.

The AA-BC technique maximizes heterogeneity along each axis, ignoring that within-group heterogeneity which is not parallel with the between-group heterogeneity. The method also maximizes discontinuities in the ordination.

Results are tentative, but when the data set has high β-diversity and/or when there are no real discontinuities in vegetational or faunal variation, AA-BC is superior to SS-BC. In theory, if a data set is rather homogeneous, association analysis will not work so well, and SS-BC may be better, although none of my data sets was homogeneous enough to show this.

In comparing the results with the best single-sample reference point method, the variance–regression criterion, the AA-BC was generally better, much better whenever heterogeneity was high or discontinuity occurred in the data set. To a long altitudinal gradient, AA-BC gave much less curvature than variance–regression BC, but several data sets were nearly identical. Perhaps some of the power of AA-BC is that it uses presence/absence data to set up the framework for the ordination, but can use quantitative data to locate the individual samples.

I also compared SS-BC with PCA ordination, since both involve a maximization of variance. The SS-BC was much better than PCA. The reason is that with the former, the variance maximized along each axis is only that associated with the between-group variance, while the variance in other directions within the group is totally ignored. This is similar to discriminant analysis, but distances still reflect species in common between groups, and correlated species are not reduced in individual importance.

The use of group centroids reduces the computation of the distance matrix, since only distances among the centroids and between centroids and individual samples need be calculated. If association analysis is used, this ordination can even be carried out without a computer, because AA is amenable to hand keysort card techniques. But, the use of group centroids in a BC ordination does put some other constraints on the method. One must decide beforehand how many ordination axes are wanted, to know how many sample groups to produce. Or, one can produce as many groups as needed to reach a certain level of homogeneity, and let that dictate how many axes to produce.

In any case, if after constructing a two-dimensional ordination, one decides a third axis is needed, one must return to the association analysis, divide the samples into more groups, and run the first two axes again and then the third. Those first two axes may be very different from the two axes of the two-dimensional ordination. In this sense, AA-BC and SS-BC are similar to the stress-minimization techniques mentioned earlier, for which one must also choose the dimension number in advance.

In summary, there are many ways to choose reference points for a BC ordination. I recommend as generally the best method an association analysis to divide the samples into relatively homogeneous groups, and then the use of those group centroids as reference points. If the data set is relatively homogeneous and continuous, or if it is more convenient to add axes onto previously constructed axes, the use of single samples is valid. I recommend as the best criterion we have for selecting such reference points the variance–regression method discussed above.

F. The Construction of Axes

Once a distance measure is decided on and two reference points are selected, the next decision is how to construct the axis. In general, it is at this point that users of BC translate a non-Euclidean (city-block or nonmetric) space into a Euclidean representation—the vectors which are the axes of an ordination. It is possible to translate the non-Euclidean distance matrix into a Euclidean space earlier, by using the Gower technique or the Orloci perpendicular axes to the $(N-1)$th dimension. It is also possible to delay the translation until after the axes are calculated. The point is that sooner or later in the process of ordination this translation is made, knowingly or unknowingly, and it is accompanied by some degree of distortion. Other alternative projections of city-block space into Euclidean representations will be described in another paper. Here I assume that the translation is accomplished in the calculation of the axes; this has been true of almost all examples of BC ordination I know, with one unintentional exception (Maycock and Curtis, 1960).

1. The First Axis

The equation for projecting sample *i* onto axis *x*, defined by two reference points A and B is (Beals, 1960)

$$x_i = (D_{AB}^2 + D_{Ai}^2 - D_{Bi}^2)/2D_{AB}$$

This equation sets the origin at reference point A. If desired the reference point may be set at the midpoint by the equation (van der Maarel, 1969):

$$x_i = (D_{Ai}^2 - D_{Bi}^2)/2D_{AB}$$

These equations treat the distances as if they were Euclidean and project a sample on a perpendicular from its true relation to reference points A and B to the line between those samples in Euclidean space. If the system is metric, there is no distortion by stretching distances beyond their real values. There will be the obvious distortion of compressing most distances except the component vector parallel to the differences in the two reference samples. If the distance matrix is nonmetric, the equation still works well. Although Orloci (1974) claims that the method requires manipulations of real Euclidean triangles to determine coordinates of the samples, this is not so. When triangularity is violated (i.e., when $a + b < c$ for three distances between any three points), the height of the triangle is imaginary; the hypercollapsed triangle may not exist in Euclidean space, but it is algebraically valid. The above equation locates such a point along the axis to minimize its "distance" in complex-number space. Thus, it minimizes the distortion in the same way it does when $a + b > c$.

Maycock and Curtis (1960), whose analysis was done before the above equations were developed, did not wish to use the time-consuming graphic method originally proposed by Bray and Curtis, and so they used the equation

$$x_i = (D_{AB} + D_{Ai} - D_{Bi})/2$$

This simply averages the two locations of sample *i* along the axis as projected from reference points A and B independently. It is also the location of the sample by city-block geometry rules (in which in a right triangle $a + b = c$). Their translation to Euclidean space was made in graphing the resultant axes. However, samples, then, are not located on the axis as close to their real position in the Euclidean graph as they are with the previous equations, because samples tend to clump toward the center of the ordination. Therefore, the earlier equations are much preferred.

Gauch and Scruggs (1979) introduced without comment three other axis construction equations, based on "one-ended proportionality," "two-ended proportionality," and "squared proportionality." These do not appear to have any geometric or nonmathematical rationale and performed less well than the geometric equation above.

2. Subsequent Axes

The second and later axes presented problems to early workers. Bray and Curtis (1957) and Beals (1960, 1965a) tried various solutions. These methods did allow some deviation of the second and later axes from perpendicularity, which Orloci (1966, 1973, 1974, 1975) continued to disagree with, despite the fact that a perpendicularizing equation was presented in 1965 (Beals, 1965b) and has been used by others (e.g., Emlen, 1972; Lechowicz and Adams, 1974). The amount of distortion due to a slightly nonperpendicular axis, however, is quite small, judged from our studies. For one real data set, a deviation of 5° for the second axis produced an efficiency 98.7% that of a perpendicularized axis, and, for another, a deviation of 10° produced 94.5% efficiency. (Efficiency is here defined as correlation squared between original and ordination distances.)

In any case, the question of distortion by nonperpendicular axes has been a moot point at Wisconsin since 1970, when we began calculating the residual distances (RD) after each axis, as mentioned in Beals (1973):

$$\mathrm{RD}_{ij} = (D_{ij}^2 - \mathrm{OD}_{ij}^2)^{1/2}$$

D_{ij} is the original distance between samples i and j and OD_{ij} is their distance across all the previous axes of the ordination. This is the distance unaccounted for by previous axes, and the matrix of these residual distances is used to construct each succeeding axis.

The methodology for each axis is therefore exactly the same as it was for the first axis, unlike the earlier BC ordination applications (through Beals, 1969a) and unlike current usage with Cornell's ORDIFLEX computer package. The use of residual distances, because it removes all distance accounted for by previous axes, necessarily results in all axes being perpendicular to one another. Provided one has a computer, this is obviously the best route to constructing succeeding axes.

Note, however, that Euclidean rules are used to calculate residual distances just as they are used to position samples along the axis. The calculation of residuals within a city-block system is complicated and will be described elsewhere. The assumption here is that the translation to Euclidean space occurs with calculation of axes. After the first axis, which may be based on a true but non-Euclidean metric, the residual distances derived by the above equation will not be metric. Violations of triangularity can occur. If reference points are far apart relative to the distance measures, such violations and consequent imaginary values in the residual matrix will be small. They should be set to zero, and they represent a slight stretching of distances in the ordination, just as do negative eigenvalues in a Gower ordination. As stated before, this distortion seems to be less harmful than the distortion induced by the use of a Euclidean metric in the first place.

VI. SURVEY OF REAL DATA PERFORMANCE

In this section is summarized the many comparative studies, formal and informal, done here at Wisconsin. In this way they may present a clearer overall picture than had they been interspersed throughout the article, although general reference was made to many of them earlier. I include in the scores only sociological studies, those using animal or plant species in samples, or resource-partitioning data. Each study enumerated below represents a distinct data set; in all, 44 such studies were consulted.

We have not made a systematic study of all ordination techniques, all distance measures, and all standardizations, and there are gaps in the analyses, but the combined results of various theses, class reports, and trial analyses (not necessarily incorporated into formal presentations), made by graduate students, colleagues, and me, are sufficient to give a clear if incomplete picture of ordination methods.

Interpretations of results are unfortunately subject to possible biases; all evaluations of ecological interpretability are subjective, and many of the reports were prepared for me by students who knew my particular viewpoint. However, students were willing to contradict me when their results suggested I was wrong, and in fact my views have changed somewhat over the past years because of them. I do not think these potential biases seriously distort the picture, but we would welcome other comparative studies from other researchers.

I have records of 30 studies which compared PCA and some form of BC ordination (mostly variance-regression endpoints). All but four clearly gave better results with BC ordination. Three of the exceptions (including Bartell *et al.*, 1978) analyzed seasonal changes in samples, and in these studies BC ordination and PCA were equally good. The fourth study actually gave better results with PCA; it consisted of only eight samples, so that the choice of vectors was extremely limited with BC ordination.

There are 11 studies comparing RA with variance–regression BC. For a first axis only, six showed a clearer axis with RA, and five with BC ordination. Considering two axes, however, all but one showed clearer environmental gradients with BC ordination. In the only comparison of AA-BC and RA, the two axes are more clearly correlated with known environmental gradients using AA-BC.

We have only five studies which have used DCA. One study showed slightly better results than using variance–regression-BC (although some discontinuities were obscured); another study suggests the two were about equal in interpretability although the ordinations were somewhat different. Two studies indicate that DCA is somewhat inferior to BC ordination, while another study found DCA to be much worse than BC ordination, worse even than RA on second and third axes. In two instances, the third axis of the DCA was more interpretable than the second axis, and that third axis was similar to the second axis of BC ordination. No comparisons with AA-BC have been made yet.

Sixteen studies compared maximum-range-BC (the original form) with variance–regression-BC: 12 found the latter clearer, two were very similar, and two others were identical. Four studies compared maximum-variance-BC (trying all pairs of samples and using the axis with greatest variance) with variance–regression-BC: the latter yielded clearer results in all cases. Maximizing variance within multidimensional-pronged clusters is simply not the best way to get ecological information. Only two studies, using Euclidean distance, looked at minimum-stress-BC vs variance–regression-BC: both showed very similar results.

Four comparisons of AA-BC and SS-BC (synthetic reference points by association analysis and by sums of squares clustering) showed that in two cases AA was better, and in two cases there was no detectable difference. Of four comparisons of AA-BC with variance–regression-BC, the former was clearer for three, and they were about equal for the fourth. The last was a very continuous data set with low β-diversity.

No Gower ordination program was available. However, the method maximizes variance along axes, and, as indicated above, that criterion has proved less useful than other criteria for axis selection, so it is doubtful if Gower ordination would be an improvement.

Unfortunately, the only stress minimization program available was Kruskal's ranking procedure, and only seven studies have used it, mainly because of constraints on computer budgets. It is a rather expensive program, and it generally needs to be run more than once on any one data set. Of seven comparisons of Kruskal ranking with variance–regression-BC, Kruskal ranking always gave clearer gradients, although in three cases it did reduce discontinuities; in the other two cases no discontinuities were apparent in either method. In four cases, several runs on the same data led to different results, indicating that local minima are a real problem. Kruskal ranking was compared with AA-BC in four studies, and the overall patterns were remarkably similar in all cases. However, Kruskal obscured discontinuities apparent in all four AA-BC ordinations.

Distance measures have also been compared. Twenty-two studies compared BC ordination using Euclidean vs Sorensen distance, and in every case the latter was superior. Four studies used the sociological favorability index to straighten and extend the sociological distance. Using Sorensen distance, comparisons between the use of that index and of relative density data suggest that there is very little improvement in the first axis, but a moderate improvement of the clarity of the second axis. None of those data sets was extreme in their β-diversity, however. The only study to compare other distance extenders was previously mentioned, in which stepping-stone distance outperformed second-order distance. Without extenders, the distances from that data set produced a very contorted elevational gradient.

Twenty studies compared presence/absence vs quantitative data in constructing BC ordinations with Sorensen distance. In eight of them presence/absence was better, in three quantitative was better, in three both gave very similar patterns, and in six the two gave different but equally interpretable patterns. The first eight were the most heterogeneous data sets, judged from the proportion of maximum distances in the distance matrix.

Of ten comparisons of relativized vs raw quantitative data, seven showed virtually no difference in the ordinations they produced, and three did; in all of these studies, the first axis for raw data in both cases represented a gradient of vegetation quantity, and the second axis reflected somewhat the first axis using relativized data. Three comparisons of relativized presence/absence data vs raw presence/absence data showed no differences, although I suspect other data sets might show one.

In three cases of plankton data and one of breeding bird communities, use of log transform improved clarity compared with raw quantitative data, but in four other, non-plankton studies, log transform showed no improvement. However, in only one of six studies in which log transform was compared with presence/absence data did the former give better results.

In ranking the ordination techniques based on results from the above, AA-BC would win handily, followed by Kruskal ranking and SS-BC. Variance–regression-BC would be next, with DCA and RA close behind. Maximum-range-BC would follow. Near the bottom would come PCA. The picture is somewhat tentative, and excludes possible entrants such as quadratic loss function and catenation, which might very well be near the top, and Gower ordination, which would certainly give better results than PCA. Of course different data sets, especially in relation to their β-diversity, may need to be treated differently.

VII. CONCLUSIONS

Ordination (in its usual sense) is intended to be a graphical and algebraical representation of major axes of compositional variation or of resource use variation. Given this intent, the Bray–Curtis technique of ordination

(BC), with the appropriate distance measure, and with either variance–regression endpoint selection or synthetic reference points, is one of the most successful and appropriate means of multivariate analysis of phytosociological and similar ecological data, and is not yet obsolete.

The most successful distance measure is consistently the one originally used, based on the Sorensen coefficient of similarity. Despite its potential for severe distortion, in practice it distorts less than Euclidean and other distances do. Both presence/absence data and relativized data may give meaningful results, the former when β-diversity is especially high, the latter when it is low. Other data adjustments may be required under special circumstances.

The most interpretable and clearest results are generally obtained when the reference points are the centroids of sample groups obtained by association analysis. These maximize patterns due to interspecific heterogeneity. If individual samples are desired as reference points, the variance–regression criterion gives the most satisfactory results. This method eliminates the problem of oddball samples, which often plagues BC ordination, and locates the longest linear axis in an often complex cluster of points.

Empirically and theoretically, principal component analysis and most other eigenvector techniques are totally inadequate unless β-diversity is exceedingly low. More serious competitors are reciprocal averaging (RA), detrended correspondence analysis (DCA), and the minimum-stress approach.

RA is not as successful in generating second and later axes as BC ordination, but it is especially effective for a single axis of a system with a single major environmental gradient. Ideally, RA should reflect species centroids and samples as points in multidimensional environmental space, but so far it is limited to only one dimension in this regard. A detrended version (DCA) corrects some weaknesses of RA, but not all. Its arbitrary and excessive manipulation of the data set makes it rather unsatisfactory, unless empirically it proves with further testing to be considerably better than the best BC methods.

The stress minimization techniques—Kruskal's nonmetric scaling, Anderson's quadratic loss function, Noy-Meir's catenation, etc.—show considerable promise, and more comparisons between them, especially the metric ones, and BC ordination are needed. But they must show marked superiority over BC ordination before they can justify replacing it, because they involve a greatly increased computational load.

ACKNOWLEDGMENTS

This article owes much to the stimulating ideas, suggestions, and arguments of many people, both faculty and students. Among the former I acknowledge especially the input of T. F. H. Allen and R. R. Kowal, although neither agrees with all that I have written. More than 60 graduate students had some role in the development of this article. Among those with substantial input to the ideas presented were S. M. Bartell, R. W. Howe, M. J. Lechowicz, W. M. Post, III, S. E. Riechert, M. R. Stromberg, and S. S. Will-Wolf. The study has been supported in part by a grant from the University of Wisconsin Graduate School.

REFERENCES

Allen, T. F. H. (1971). Multivariate approaches to the ecology of algae on terrestrial rock surfaces in North Wales. *J. Ecol.* **59**, 803–826.

Allen, T. F. H. (1981). The signal in ordination curvature: structure between levels of organization (unpublished manuscript).

Allen, T. F. H., and Koonce, J. F. (1973). Multivariate approaches to algal strategems and tactics in systems analysis of phytoplankton. *Ecology* 54, 1234–1246.

Anderson, A. J. B. (1971). Ordination methods in ecology. *J. Ecol.* **59**, 713–726.

Austin, M. P. (1968). An ordination study of a chalk grassland community. *J. Ecol.* **56**, 739–757.

Austin, M. P. (1976a). On non-linear species response models in ordination. *Vegetatio* **33**, 33–41.

Austin, M. P. (1976b). Performance of four ordination techniques assuming three different non-linear species response models. *Vegetatio* **33**, 43–49.

Austin, M. P. (1977). Use of ordination and other multivariate descriptive methods to study succession. *Vegetatio* **35**, 165–175.

Austin, M. P., and Greig-Smith, P. (1968). The application of quantitative methods to vegetation survey. II. Some methodological problems of data from rain forest. *J. Ecol.* **56**, 827–844.

Austin, M. P., and Noy-Meir, I. (1971). The problem of non-linearity in ordination: Experiments with two-gradient models. *J. Ecol.* **59**, 763–773.

Austin, M. P., and Orloci, L. (1966). Geometric models in ecology. II. An evaluation of some ordination techniques. *J. Ecol.* **54**, 217–227.

Bakuzis, E. V., and Hansen, H. L. (1959). A provisional assessment of species synecological requirements in Minnesota forests. *Minn. For. Notes* **84**.

Bannister, P. (1968). An evaluation of some procedures used in simple ordinations. *J. Ecol.* **56**, 27–34.

Barkham, J. P., and Norris, J. M. (1970). Multivariate procedures in an investigation of vegetation and soil relations of two beech woodlands, Cotswold Hills, England. *Ecology* **51**, 630–639.

Bartell, S. M. (1973). A multivariate statistical approach to the phytoplankton community structure and dynamics in the Lake Wingra ecosystem. M.Sc. Thesis, University of Wisconsin, Madison.

Bartell, S. M., Allen, T. F. H., and Koonce, J. F. (1978). An assessment of principal component analysis for description of phytoplankton periodicity in Lake Wingra. *Phycologia* **17**, 1–11.

Beals, E. W. (1960). Forest bird communities in the Apostle Islands of Wisconsin. *Wilson Bull.* **72**, 156–181.

Beals, E. W. (1965a). Ordination of some corticolous communities in south-central Wisconsin. *Oikos* **16**, 1–8.

Beals, E. W. (1965b). Species patterns in a Lebanese Poterietum. *Vegetatio* **13**, 69–87.

Beals, E. W. (1969a). Vegetation of the Erer-Gota Plain, Ethiopia. *J. Ecol.* **57**, 655–667.

Beals, E. W. (1969b). Vegetational change along altitudinal gradients. *Science* **165**, 981–985.

Beals, E. W. (1973). Ordination: Mathematical elegance and ecological naivete. *J. Ecol.* **61**, 23–35.

Beals, E. W., and Cottam, G. (1960). The forest vegetation of the Apostle Islands, Wisconsin. *Ecology* **41**, 785–790.

Benzecri, J. P. (1964). Sur l'analyse factorielle des proximites. *Publ. Inst. Stat., Univ. Paris* **13**, 235–282.

Bouxin, G. (1976). Ordination and classification in the upland Rugege Forest (Rwanda, Central Africa). *Vegetatio* **32**, 97–115.

Boyd, V. T., and Allen, T. F. H. (1981). Value ascribed to objects: Multivariate elucidation of a conceptual structure. *J. Soc. Biol. Struct.* **4**, 39–57.

Bray, J. R., and Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27**, 325–349.

Carey, A. B., McLean, R. G., and Maupin, G. O. (1980). The structure of a Colorado tick fever ecosystem. *Ecol. Monogr.* **50**, 131–151.

Carleton, T. J. (1980). Non-centered component analysis of vegetation data: a comparison of orthogonal and oblique rotation. *Vegetatio* **42**, 59–66.

Cassie, R. M. (1969). Multivariate analysis in ecology. *Proc. N. Z. Ecol. Soc.* **16**, 53–57.

Cassie, R. M., and Michael, A. D. (1968). Fauna and sediments of an intertidal mud flat: A multivariate analysis. *J. Exp. Mar. Biol. Ecol.* **2**, 1–23.

Christensen, N. L., and Peet, R. K. (1981). Secondary forest succession on the North Carolina Piedmont. *In* "Forest succession: Concepts and applications" (D. C. West, H. H. Shugart, and D. B. Botkin, eds.), pp. 230–245. Springer-Verlag, Berlin and New York.

Clymo, R. (1980). Preliminary survey of the peat-bog Hummell Knowe Moss using various numerical methods. *Vegetatio* **42**, 129–148.

Colburn, E. A. (1975). Distribution and ecology of aquatic pulmonate snails on a Wisconsin River floodplain. M.Sc. Thesis, University of Wisconsin, Madison.

Conner, R. H., and Adkinsson, C. S. (1977). Principal component analysis of woodpecker nesting habitat. *Wilson Bull.* **89**, 122–129.

Cottam, G., Goff, F. G., and Whittaker, R. H. (1973). Wisconsin comparative ordination. *Handb. Veg. Sci.* **5**, 193–221.

Curtis, J. T. (1959). "The Vegetation of Wisconsin. An Ordination of Plant Communities." University of Wisconsin Press, Madison.

Curtis, J. T., and McIntosh, R. P. (1951). An upland forest continuum of the prairie-forest border region of Wisconsin. *Ecology* **32**, 476–496.

Dagnelie, P. (1960). Contribution a l'etude des communautes vegetales par l'analyse factorielle. *Bull. Serv. Carte Phytogeog. Ser. B* **5**, 7–71, 93–195.

Dagnelie, P. (1973). L'analyse factorielle. *Handb. Veg. Sci.* **5**, 223–248.

Dale, M. B. (1975). On objectives of methods of ordination. *Vegetatio* **30**, 15–32.

DeJong, M. D. (1976). The distribution of breeding birds in relation to vegetation in lowland forests of southern Wisconsin. M.Sc. Thesis, University of Wisconsin, Madison.

del Moral, R. (1980). On selecting indirect ordination methods. *Vegetatio* **42**, 75–84.

del Moral, R., and Watson, A. F. (1978). Gradient structure of forest vegetation in the central Washington Cascades. *Vegetatio* **38**, 29–48.

Denslow, J. S. (1977). Secondary succession in a Colombian rainforest: Strategies of species response along a disturbance gradient. Ph.D. Thesis, University of Wisconsin, Madison.

Dueser, R. D., and Shugart, H. H., Jr. (1979). Niche pattern in a forest-floor small-mammal fauna. *Ecology* **60**, 108–118.

Emlen, J. T. (1972). Size and structure of a wintering avian community in southern Texas. *Ecology* **53**, 317–329.

Emlen, J. T. (1977). Land bird communities of Grand Bahama Island: The structure and dynamics of an avifauna. *Ornithol. Monogr.* **24**, 1–129.

Fasham, M. J. R. (1977). A comparison of non-metric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes. *Ecology* **58**, 551–561.

Feoli, E. (1977). A criterion for monothetic classification of phytosociological entities on the basis of species ordination. *Vegetatio* **33**, 147–152.

Futuyma, D. J., and Gould, F. (1979). Associations of plants and insects in a deciduous forest. *Ecol. Monogr.* **49**, 33–50.

Gauch, H. G., Jr. (1973a). The relationship between sample similarity and ecological distance. *Ecology* **54**, 618–622.

Gauch, H. G., Jr. (1973b). A quantitative evaluation of the Bray-Curtis ordination. *Ecology* **54**, 829–836.

Gauch, H. G., Jr. (1977). ORDIFLEX: A flexible computer program for four ordination techniques. Cornell University, Department of Ecology and Systematics.

Gauch, H. G., Jr. (1982). "Multivariate analysis in community ecology." Cambridge Univ. Press, London and New York.

Gauch, H. G., Jr., and Scruggs, W. M. (1979). Variants of polar ordination. *Vegetatio* **40**, 147–153.

Gauch, H. G., Jr., and Wentworth, T. R. (1976). Canonical correlation analysis as an ordination technique. *Vegetatio* **33**, 17–22.

Gauch, H. G., Jr., and Whittaker, R. H. (1972). Comparison of ordination techniques. *Ecology* **53**, 868–875.

Gauch, H. G., Jr., Chase, G. B., and Whittaker, R. H. (1974). Ordination of vegetation samples by Gaussian species distributions. *Ecology* **55**, 1382–1390.

Gauch, H. G., Jr., Whittaker, R. H., and Wentworth, T. R. (1977). A comparative study of reciprocal averaging and other ordination techniques. *J. Ecol.* **65**, 157–174.

Gittins, R. (1965). Multivariate approaches to a limestone grassland community. II. A direct species ordination. *J. Ecol.* **53**, 403–409.

Goff, F. G., and Cottam, G. (1967). Gradient analysis: The use of species and synthetic indices. *Ecology* **48**, 793–806.

Goldstein, R. A., and Grigal, D. F. (1972). Definition of vegetation structure by canonical analysis. *J. Ecol.* **60**, 277–284.

Goodall, D. W. (1954). Objective methods for the classification of vegetation III. An essay in the use of factor analysis. *Aust. J. Bot.* **2**, 304–324.

Goodall, D. W. (1970). Statistical plant ecology. *Annu. Rev. Ecol. Syst.* **1**, 99–124.

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325–338.

Green, R. H. (1971). A multivariate statistical approach to the Hutchinsonian niche: Bivalve molluscs of central Canada. *Ecology* **52**, 543–556.

Green, R. H. (1979). "Sampling Design and Statistical Methods for Environmental Biologists." Wiley (Interscience), New York.

Greig-Smith, P. (1964). "Quantitative Plant Ecology," 2nd ed. Butterworth, London.

Grigal, D. F., and Goldstein, R. A. (1971). An integrated ordination-classification analysis of an intensively sampled oak-hickory forest. *J. Ecol.* **59**, 481–492.

Hall, J. B., and Swaine, M. D. (1976). Classification and ecology of closed-canopy forest in Ghana. *J. Ecol.* **64**, 913–951.

Hanski, I., and Koskela, H. (1977). Niche relations among dung-inhabiting beetles. *Oecologia* **28**, 203–231.

Hatheway, W. H. (1971). Contingency-table analysis of rain forest vegetation. *In* "Statistical Ecology," (G. P. Patil, E. C. Pielou, and W. E. Waters, eds.), Vol. 3, pp. 271–313. Penn. State Univ. Press, University Park.

Hill, M. O. (1973). Reciprocal averaging: An eigenvector method of ordination. *J. Ecol.* **61**, 237–249.

Hill, M. O. (1979). DECORANA, A fortran program for detrended correspondence analysis and reciprocal averaging. Cornell University, Department of Ecology and Systematics, Ithaca, New York.

Hill, M. O., and Gauch, H. G., Jr. (1980). Detrended correspondence analysis: An improved ordination technique. *Vegetatio* **42**, 47–58.

Holmes, R. T., Bonney, R. E., Jr., and Pacala, S. W. (1979). Guild structure of the Hubbard Brook bird community: A multivariate approach. *Ecology* **60**, 512–520.

Howe, R. W. (1977). An ecological analysis of small-island bird communities in northern Minnesota. M.Sc. Thesis, University of Wisconsin, Madison.

Ihm, P., and van Groenewoud, H. (1975). A multivariate ordering of vegetation data based on Gaussian type gradient response curves. *J. Ecol.* **63**, 767–777.

Ivimey-Cook, R. B., and Proctor, M. C. F. (1967). Factor analysis of data from an East Devon heath: A comparison of principal components and rotated solutions. *J. Ecol.* **55**, 405–413.

James, F. C. (1971). Ordinations of habitat relationships among breeding birds. *Wilson Bull.* **83**, 215–236.

Jeglum, J. K., Wehrhann, C. P., and Swan, J. M. A. (1971). Comparisons of environmental ordinations with principal component vegetational ordinations for sets of data having differnt degrees of complexity. *Can. J. For. Res.* **1**, 99–122.

Jesberger, J. A., and Sheard, J. W. (1973). A quantitative study and multivariate analysis of corticolous lichen communities in the southern boreal forest of Saskatchewan. *Can. J. Bot.* **51**, 185–201.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* **23**, 187–200.

Kantak, G. E. (1977). Partitioning of fruit resources by some birds of the Yucatan Peninsula, Mexico. M.Sc. Thesis, University of Wisconsin, Madison.

Karr, J. R., and James, F. C. (1975). Ecomorphological configurations and convergent evolution in species and communities. *In* "Ecology and Evolution of Communities" (M. L. Cody and J. M. Diamond, eds.), pp. 258–291. Harvard Univ. Press, Cambridge, Massachusetts.

Kelsey, C. T., Goff, F. G., and Fields, D. (1977). Theory and analysis of vegetation pattern. *Environ. Sci. Div. Publ.* **922**, Oak Ridge National Laboratory.

Kercher, J. R., and Goldstein, R. A. (1977). Analysis of an east Tennessee oak hickory forest by canonical correlation of species and environmental parameters. *Vegetatio* **35**, 153–163.

Kessell, S. R., and Whittaker, R. H. (1976). Comparisons of three ordination techniques. *Vegetatio* **32**, 21–29.

Kline, V. M. (1976). Dynamics of the vegetation of a small watershed. Ph.D. Thesis, University of Wisconsin, Madison.

Knight, D. H. (1965). A gradient analysis of Wisconsin prairie vegetation on the basis of plant structure and function. *Ecology* **46**, 744–747.

Knight, D. H., and Loucks, O. L. (1969). A quantitative analysis of Wisconsin forest vegetation on the basis of plant function and gross morphology. *Ecology* **50**, 219–234.

Kruskal, J. B. (1964). Non-metric multidimensional scaling: A numerical method. *Psychometrika* **29**, 115–129.

Lambert, J. M., and Dale, M. B. (1964). The use of statistics in phytosociology. *Adv. Ecol. Res.* **2**, 59–99.

Lechowicz, M. J., and Adams, M. S. (1974). Ecology of *Cladonia* lichens. I. Preliminary assessment of the ecology of terricolous lichen-moss communities in Ontario and Wisconsin. *Can. J. Bot.* **52**, 55–64.

Loucks, O. L. (1962). Ordinating forest communities by means of environmental scalars and phytosociological indices. *Ecol. Monogr.* **32,** 137–166.

McIntosh, R. P., and Hurley, R. T. (1964). The spruce-fir forests of the Catskill mountains. *Ecology* **45,** 314–326.

Marks, P. L., and Harcombe, P. A. (1981). Forest vegetation of the big thicket, southeast Texas. *Ecol. Monogr.* **51,** 287–305.

Matthews, J. A. (1978). An application of non-metric multidimensional scaling to the construction of an improved species plexus. *J. Ecol.* **66,** 157–173.

Matthews, J. A. (1979). A study of the variability of some successional and climax plant assemblage-types using multiple discriminant analysis. *J. Ecol.* **67,** 255–271.

Maycock, P. F., and Curtis, J. T. (1960). The phytosociology of boreal conifer-hardwood forests of the Great Lakes region. *Ecol. Monogr.* **30,** 1–35.

Mohler, C. L. (1981). Effects of sample distribution along gradients of eigenvector ordination. *Vegetatio* **45,** 141–145.

Monk, C. D. (1965). Southern mixed hardwood forests of northcentral Florida. *Ecol. Monogr.* **35,** 335–354.

Mowbray, T. B., and Oosting, H. J. (1968). Vegetation gradients in relation to environment and phenology in a southern Blue Ridge gorge. *Ecol. Monogr.* **38,** 309–344.

Newsome, R. D., and Dix, R. L. (1968). The forests of the Cypress Hills, Alberta and Saskatchewan, Canada. *Am. Midl. Nat.* **80,** 118–185.

Nichols, S. (1977). On the interpretation of principal components analysis in ecological contexts. *Vegetatio* **34,** 191–197.

Noon, B. R. (1981). The distribution of an avian guild along a temperate elevational gradient: The importance and expression of competition. *Ecol. Monogr.* **51,** 105–124.

Norris, J. M., and Barkham, J. P. (1970). A comparison of some Cotswold beechwoods using multiple-discriminant analysis. *J. Ecol.* **58,** 603–619.

Noy-Meir, I. (1971). Multivariate analysis of the semi-arid vegetation in south-eastern Australia: Nodal ordination by component analysis. *Proc. Ecol. Soc. Aust.* **6,** 159–193.

Noy-Meir, I. (1973). Data transformations in ecological ordination. I. Some advantages of non-centering. *J. Ecol.* **61,** 329–341.

Noy-Meir, I. (1974). Catenation: Quantitative methods for the definition of coenoclines. *Vegetatio* **29,** 89–99.

Noy-Meir, I., and Whittaker, R. H. (1977). Continuous multivariate methods in community analysis: Some problems and developments. *Vegetatio* **33,** 79–98.

Noy-Meir, I., Walker, D., and Williams, W. T. (1975). Data transformations in ecological ordination. II. On the meaning of data standardization. *J. Ecol.* **63,** 779–800.

Orloci, L. (1966). Geometric models in ecology. I. The theory and application of some ordination methods. *J. Ecol.* **54,** 193–215.

Orloci, L. (1967). An agglomerative method for classification of plant communities. *J. Ecol.* **55,** 193–206.

Orloci, L. (1973). Ordination by resemblance matrices. *Handb. Veg. Sci.* **5,** 249–286.

Orloci, L. (1974). Revisions for the Bray and Curtis ordination. *Can. J. Bot.* **52,** 1773–1776.

Orloci, L. (1975). "Multivariate Analysis in Vegetation Research." Junk, The Hague.

Pemadasa, M. A., and Mueller-Dombois, D. (1979). An ordination study of montane grasslands of Sri Lanka. *J. Ecol.* **67,** 1009–1023.

Persson, S. (1981). Ecological indicator values as an aid in the interpretation of ordination diagrams. *J. Ecol.* **69,** 71–84.

Phillips, D. L. (1978). Polynomial ordination: Field and computer simulation testing of a new method. *Vegetatio* **31,** 129–140.

Pielou, E. C. (1977). "Mathematical Ecology." Wiley (Interscience), New York.
Prentice, I. C. (1977). Non-metric ordination methods in ecology. *J. Ecol.* **68,** 85–94.
Prentice, I. C. (1980). Vegetation analysis and order invariant gradient models. *Vegetatio* **42,** 27–34.
Ramensky, L. G. (1930). Zur Methodik der vergleichenden Bearbeitung und Ordnung von Pflanzenlisten und anderen Objekten, die durch mehrere verscheidenartig wirkende Fakturen bestimmt werden. *Beitr. Biol. Pflanz.* **18,** 269–304.
Ricklefs, R. E., and Travis, J. (1980). A morphological approach to the study of avian community organization. *Auk* **97,** 321–338.
Risser, P. G., and Rice, E. L. (1971). Phytosociological analysis of Oklahoma upland forest species. *Ecology* **52,** 940–945.
Robertson, P. A. (1978). Comparison of techniques for ordinating and classifying old-growth floodplain forests in southern Illinois. *Vegetatio* **37,** 43–51.
Rusterholz, K. A. (1973). Island bird communities on Burntside Lake, Minnesota. M.Sc. Thesis, University of Wisconsin, Madison.
Rusterholz, K. A. (1979). Niche relations of pine foliage-gleaning birds in different competitive regimes. Ph.D. Thesis, University of Wisconsin, Madison.
Sabo, S. R. (1980). Niche and habitat relations in subalpine bird communities of the White Mountains of New Hampshire. *Ecol. Monogr.* **50,** 241–259.
Sabo, S. R., and Whittaker, R. H. (1979). Bird niches in a subalpine forest: An indirect ordination. *Proc. Natl. Acad. Sci. U.S.A.* **76,** 1338–1342.
Sanford, W. W. (1974). The use of epiphytic orchids to characterize vegetation in Nigeria. *Bot. J. Linn. Soc.* **68,** 291–301.
Schnell, G. D., Risser, P. G., and Helsel, J. F. (1977). Factor analysis of tree distribution patterns in Oklahoma. *Ecology* **58,** 1345–1355.
Shepard, R. N., and Carroll, J. D. (1966). Parametric representation of non-linear data structures. In "Multivariate Analysis" (P. R. Krishnaiah, ed.), pp. 561–592. Academic Press, New York.
Shepherd, J. D. (1975). The phytosociology of a tropical seasonal forest in southeastern Mexico. M.Sc. Thesis, University of Wisconsin, Madison.
Sibson, R. (1972). Order invariant methods for data analysis. *J. R. Stat. Soc. Ser. B* **34,** 311–349.
Smartt, P. F. M., Meacock, S. E., and Lambert, J. M. (1974). Investigations into the properties of quantitative vegetational data. I. Pilot study. *J. Ecol.* **62,** 735–759.
Smartt, P. F. M., Meacock, S. E., and Lambert, J. M. (1976). Investigations into the properties of quantitative vegetational data. II. Further data type comparisons. *J. Ecol.* **64,** 41–78.
Stephenson, G. R. (1974). Social structure of mating activity in Japanese macaques. *Proc. Symp. Int. Primatol. Soc. Congr. 5th, 1974,* pp. 63–115.
Swan, J. M. A. (1970). An examination of some ordination problems by use of simulated vegetational data. *Ecology* **51,** 89–102.
Swan, J. M. A., and Dix, R. L. (1966). The phytosociological structure of upland forest at Candle Lake, Saskatchewan. *J. Ecol.* **54,** 13–40.
Swan, J. M. A., Dix, R. L. and Wehrhahn, C. F. (1969). An ordination technique based on the best possible stand-defined axes and its application to vegetational analysis. *Ecology* **50,** 206–212.
Terborgh, J. (1971). Distribution on environmental gradients: Theory and a preliminary interpretation of distributional patterns in the avifauna of the Cordillera Vilcabamba, Peru. *Ecology* **52,** 23–40.
Thomson, J. D. (1975). Some community-level aspects of a bog pollination system. M.Sc. Thesis, University of Wisconsin, Madison.

Tyler, C. (1981). Geographic variation in Fennoscandian and Estonian *Schoenus* wetlands. *Vegetatio* **45**, 165–182.

van der Maarel, E. (1969). On the use of ordination models in phytosociology. *Vegetatio* **19**, 21–46.

van der Maarel, E. (1972). Ordination of plant communities on the basis of their plant genus, family and order relationships. *In* "Grundfragen und Methoden in der Pflanzensoziologie" (R. Tuxen, ed.), pp. 183–192. Junk, The Hague.

van Groenewoud, H. (1976). Theoretical considerations on the covariation of plant species along ecological gradients with regard to multivariate analysis. *J. Ecol.* **64**, 837–847.

Waide, R. B. (1973). Seasonal changes in the avifauna of a tropical wet forest. M.Sc. Thesis, University of Wisconsin, Madison.

Walker, B. H. (1974). Some problems arising from the preliminary manipulation of plant ecological data for subsequent numerical analysis. *J. South Afr. Bot.* **40**, 1–13.

Walker, B. H. (1975). Vegetation–site relationships in the Harvard forest. *Vegetatio* **29**, 169–178.

Westman, W. E. (1975). Edaphic climax pattern of the pygmy forest region of California. *Ecol. Monogr.* **45**, 109–135.

Whittaker, R. H. (1952). A study of summer foliage insect communities in the Great Smoky Mountains. *Ecol. Monogr.* **22**, 1–44.

Whittaker, R. H. (1956). Vegetation of the Great Smoky Mountains. *Ecol. Monogr.* **26**, 1–80.

Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.* **30**, 279–338.

Whittaker, R. H. (1967). Gradient analysis of vegetation. *Biol. Rev.* **42**, 207–264.

Whittaker, R. H., and Gauch, H. G., Jr. (1973). Evaluation of ordination techniques. *Handb. Veg. Sci.* **5**, 287–321.

Whittaker, R. H., and Gauch, H. G., Jr. (1978). Evaluation of ordination techniques. *In* "Ordination of Plant Communities" (R. H. Whittaker, ed.), pp. 277–336. Junk, The Hague.

Wiegleb, G. (1980). Some applications of principal components analysis in vegetation: Ecological research of aquatic communities. *Vegetatio* **42**, 67–73.

Williams, W. J., and Lambert, J. M. (1959). Multivariate methods in plant ecology. I. Association–analysis in plant communities. *J. Ecol.* **47**, 83–101.

Williams, W. T., Lambert, J. M., and Lance, G. B. (1966). Multivariate methods in plant ecology. V. Similarity analyses and information-analysis. *J. Ecol.* **54**, 427–445.

Williamson, M. H. (1978). The ordination of incidence data. *J. Ecol.* **66**, 911–920.

Will-Wolf, S. S. (1980). Structure of corticolous lichen communities before and after exposure to emission from a "clean" coal-fired generating station. *Bryologist* **83**, 281–295.

Wolf, S. W. (1975). Multivariate analysis of foraging site selection by flower-feeding insects in a western South Dakota prairie. Ph.D. Thesis, University of Wisconsin, Madison.

Wood, L. E. (1979). A comparison of breeding bird communities in northern Wisconsin pine forests. M.Sc. Thesis, University of Wisconsin, Madison.