

Comparison of ordinations and classifications of vegetation data

János Podani

Department of Plant Taxonomy and Ecology, L. Eötvös University, Kun B. tér 2, H-1083 Budapest, Hungary

Accepted 28.3.1989

Keywords: Consensus method, Dendrogram, Matrix correlation, Multiple comparison, Partition, Procrustes analysis, Resemblance matrix

Abstract

The methodology of comparing the results of multivariate community studies (resemblance matrices, ordinations, hierarchical and nonhierarchical classifications) is reviewed from two viewpoints: basic strategy and measure employed. The basic strategy is determined by 7 choices concerning the type of results, consensus methods or resemblance measures, hypothesis testing or exploratory analysis, lack or presence of reference basis, data set congruence or algorithmic effects, number of factors responsible for differences among results, and the number of properties considered in the comparison. Included is a brief summary of methods applicable to vegetation studies. Examples from a grassland survey demonstrate the utility of comparisons in evaluating the effects of plot size, data type, standardization, taxonomic level and number of species on classifications and ordinations.

Abbreviations: OUC = Operational Unit of Comparison; PCA = Principal Components Analysis; PCoA = Principal Coordinates Analysis; SSA = Incremental Sum of Squares Agglomeration.

Introduction

In the past two decades multivariate techniques have been extensively applied to the classification and ordination of vegetational entities (see Orlóci 1978; van der Maarel 1979; Gauch 1982; Greig-Smith 1983; Legendre & Legendre 1983; Pielou 1984; Kershaw & Looney 1985; Digby & Kempton 1987). It is commonly observed that different clustering algorithms produce more or less dissimilar classifications when applied to the same set of data, unless the underlying group structure is obvious. Similar statement holds for ordinations, which may substantially differ with the scaling technique used. As the number of

classification and ordination methods suggested for application to community studies is steadily increasing, it is essential to know the relative merits of techniques. Clearly, this problem calls for comparative studies. The evaluation of methods may primarily be done with reference to known properties of simulated data (see e.g., Bayne *et al.* 1980; Milligan & Isaac 1980 and Gauch & Whittaker 1981 for assessment of clustering algorithms; and Austin 1976; Fasham 1977; Gauch *et al.* 1977; Gauch *et al.* 1981; Orlóci *et al.* 1984; Kenkel & Orlóci 1986 and Minchin 1987 for comparisons of scaling procedures). In these cases interest is focused on the ability of multivariate techniques to reveal certain

data structures (e.g., groups of objects or non-linear gradients). This is a general problem, however, which arises not only in community analysis but in any field of science that utilizes multivariate methods.

In biology, there are many specific situations when direct comparison of outputs of alternative analyses of the same objects becomes crucial, without explicit references to the properties of data. In this respect, much of the work has been done by taxonomists. Rohlf & Sokal (1981) and Day (1983) listed several issues of numerical taxonomy in which comparison of results is essential. Taxonomic congruence, stability of classifications and character state predictivity are cases in point. Analogous problems do exist in vegetation science, because the selection of a particular multivariate technique is only one of the many decisions that have to be made during a computer-assisted survey. In addition, the sampling characteristics (plot size, shape and arrangement, etc.), data type and data transformation may also influence the final results. The analysis of these effects, just as the examination of congruence of results derived from the application of different character sets also require objective evaluation. When permanent plots are employed to study succession, the change of results over time may necessitate such work. These examples are sufficient to show that in vegetation science comparative studies are not restricted to the assessment of the performance of multivariate techniques.

This paper emphasizes the importance of direct comparisons in vegetation studies. First, I describe basic strategies by introducing a framework of choices potentially made when results of vegetation analysis are assessed. Past contributions to the topic are reviewed throughout the text. Although resemblance matrices are not mentioned in the title, the problem of their comparison is also included since the evaluation of classifications and ordinations is often reduced to matrix comparisons. Existing methodologies are summarized very briefly, however. For more information on the technical details of comparisons, the reader should consult the literature cited. The

final section of the paper presents illustrative examples based on actual data from a rock grassland survey.

Basic strategies

The principal idea is that a particular result is treated as an individual object, just like plant individuals in taxonomy or relevés in phytosociology. Since this paper is concerned with four different kinds of results, i.e., resemblance (proximity) matrices, ordinations, partitions and dendrograms, it seems useful to introduce the collective term *operational unit of comparison*, OUC, to cover all of them. OUC will be analogous to the OTUs of numerical taxonomy (Sneath & Sokal 1973) and to the OGU's of biogeography (Crovello 1981). The abbreviation OUC will be used whenever general reference is made to results.

The strategy of comparison is determined by the objectives of the study, by the limitations imposed by inherent properties of the data and by the restrictions of techniques. Therefore, one is faced with several choices to decide on the basic strategy, and more opportunities regarding a particular technique for comparison. The subsequent discussion follows a somewhat arbitrary sequence of choices and does not imply any order of importance. It is also noted that many combinations of alternatives do not make sense or are irrelevant to vegetation studies, and we do not have to make all the 7 choices in every instance.

Choice 1: inter-type vs intra-type comparison

The OUCs involved in inter-type comparisons are not of the same kind. Numerical taxonomy provides many examples for this approach. The most common combination of types of results is in the comparison of a matrix, representing a dendrogram, with a resemblance matrix from which the classification was obtained (cophenetic comparisons, cf. Rohlf & Sokal 1981; see Robertson 1979 for an example from vegetation ecology). Simi-

larly, ordinations of objects in few dimensions may be contrasted with the original resemblance matrix (e.g., Jeglum *et al.* 1971). In fact, this is the strategy followed step by step in non-metric multi-dimensional scaling (Kruskal 1964). The comparison of ordinations or classifications is possible even with the raw data (Duncan & Estabrook 1976; Gower 1983; Ferraris 1983). Ordinations and partitions are routinely contrasted with each other to find their correspondence. The comparison of a given partition with several hierarchical levels of a dendrogram is also conceivable. Measurements of how well an ordination coincides with the known configuration of objects along gradients are logically inter-type comparisons as well. At an even more general level, inter-type comparisons include any measurements of 'fit' between data and a particular pattern model.

The discussion in the present paper will be focused entirely on the other alternative, the intra-type comparisons, when all OUCs being evaluated are of the same kind. Reading through the literature I had the impression that the topic of inter-type comparisons would deserve a separate paper. It is noted, however, that some techniques (especially matrix correlation) may be equally applied in both strategies.

Choice 2: resemblance vs consensus

One may wish to express by a single number the resemblance (i.e., similarity, dissimilarity, distance, or other relationship) between two OUCs. If each OUC can be described in terms of an X vector variable, this goal is readily achieved by most resemblance coefficients known from numerical taxonomy and ecology. An example is the product-moment correlation coefficient, which is commonly employed in matrix comparisons (Rohlf 1974). There are, of course, situations in which formulae designed exclusively for the solution of special problems are used.

The other possibility is to reject the pairwise comparisons by finding a consensus OUC for $k \geq 2$ alternative OUCs. In the broadest sense, a

consensus object is defined as an OUC representing shared structural properties of the rivals, but differences may also be emphasized (see e.g., Adams 1972; Day 1988; Faith 1988, for details). Consensus generation is a fruitful strategy if the same data set is analyzed by several techniques, each considered equal in importance, and only a single OUC that best reflects the common properties of the competing results is sought. A consensus OUC may also be obtained from OUCs that are based on separate subsets of attributes. As Lefkovich (1985a) noted, this is done with the assumption that such an OUC may be closer to the 'true' relationships than that calculated from the pooled data. However, there is no proof for this statement.

The history of consensus methods is reviewed by Day (1988). The majority of methods have been developed to synthesize several evolutionary trees into a consensus tree (see Rohlf 1982; and references therein). In vegetation science, these techniques will hardly, if ever, be used, since the fine branches of dendrograms are of no practical interest. Instead, consensus partitions and ordinations seem more promising. An early attempt for constructing consensus partitions in vegetation analysis is due to Pritchard & Anderson (1971). A more sophisticated approach has been adopted by Podani (1989a) the method is applied to alternative classifications of grassland communities. Examples for the use of consensus ordinations (i.e., 'average' configurations) are given in Digby & Kempton (1987) for animal assemblages and in Kenkel & Booth (1987) for marine fungal communities.

The consensus object for a pair of OUCs may often be used to derive some similarity index for the OUCs, so there is a close relationship between the alternative choices. Rohlf (1982) and Faith (1988) listed some examples for this approach. Moreover, Faith & Belbin (1986) suggested combining resemblance measures and consensus similarity into a family of functions for the comparison of classifications.

Choice 3: hypothesis testing vs exploratory analysis

An obvious question regarding a concrete value of inter-OUC resemblance: is that particular outcome statistically significant? In other words, what is the probability for getting that outcome if the two OUCs are in fact randomly generated? The answers to these questions may have serious ecological implications. For example, Orłóci (1978) suggested the use of an information statistic for testing the independence of two partitions of relevés, the first based on floristic data and the second based on soil properties. In such cases, a significantly high resemblance indicates high predictive power of either classification relative to the other and suggests a close relationship between floristic and soil variables.

Of course, statistical tests cannot, or should not be performed without satisfying some basic conditions. Parametric techniques are the most often misused, since the assumption of normality is rarely met. Therefore, just to mention an example, the Pearson product-moment correlation coefficient calculated between the corresponding values of two distance matrices should not be tested for significance (cf. Dietz 1983). The interdependence among the off-diagonal values is another reason to disregard the parametric methods (Farris 1973). These difficulties may be overcome by permutation tests (Mantel 1967; Dietz 1983; Burgman 1987) or special nonparametric methods (Lefkovitch 1984; 1985b). For partitions, the information statistic is most appropriate (Orłóci 1978; Feoli *et al.* 1984a).

In fact, any coefficient of resemblance may be tested for significance, even if its distributional properties are unknown. Axiomatic (exact) and empirical (sampled randomization) tests (cf. Rohlf & Sokal 1981) offer a solution for this problem (examples are found in Penny *et al.* 1982; Shao & Rohlf 1983; and Podani 1986). Significance tests are invalid if the two OUCs being compared have been derived from the same data set, since the condition of independence is failed (cf. Sokal 1979; Dietz 1983).

The bootstrap method, applied first by Felsenstein (1985) to biological taxonomy, may be used

to define consensus OUCs with specified confidence limits. In cladistic studies, for example, the set of characters is resampled many times with replacement to produce new sets of the same size as original, and a tree is determined for each case. Then, groups occurring in at least, say, 95% of the results are included in the consensus tree. It remains to be investigated whether bootstrapped consensus objects are useful tools to evaluate analogous problems in community analysis.

If the conditions for hypothesis testing are suspect for any reason, the actual resemblance coefficient may still prove informative for exploratory or descriptive purposes. Clearly, if we have only two OUCs, the single resemblance value can hardly describe anything, as no good reference basis is available. However, if many OUCs are to be evaluated, the pairwise comparisons may be contrasted with one another. In this case, one is faced with yet another choice, which is discussed below.

Choice 4: overall comparison vs comparisons to a reference OUC

If the set of k OUCs to be assessed lacks a particular element that could be used as a reference, the only strategy we can follow is to calculate all the $k(k-1)/2$ pairwise resemblance coefficients. Then, the resemblance matrix of OUCs may be subjected to multivariate analysis to reveal underlying relationships among the results. Such overall comparison or 'super-taxonomy' (e.g., classification of classifications) has long served comparative purposes in taxonomy and statistics (e.g., Schnell 1970; Phipps 1971; Arabie & Boorman 1973; Dubes & Jain 1976; Duncan *et al.* 1980; Podani & Dickinson 1984; Paule & Gömöry 1987; Dickinson *et al.* 1988) and has several applications to plant community studies as well (Booth 1978; Gauch 1980; Gauch & Whittaker 1981; Feoli *et al.* 1984b; Kenkel & Booth 1987; Podani 1985, 1989b). Most new examples presented in this paper will illustrate this exploratory approach.

When spatial processes (*sensu* Podani 1984)

are evaluated in vegetation studies, there is often a standard OUC whose comparison with all the others is more meaningful than the remaining comparisons. In such cases, multivariate analysis is not necessary, as the change of OUCs may be monitored using profile diagrams. A classical example was presented by Orlóci & Mukkattu (1973). They evaluated the effect of species number upon resemblance structures. Obviously, the reference OUC was the resemblance matrix calculated using all the species present in the data. Podani (1986) provided further examples for this approach by examining the effect of sample size on community classification.

Choice 5: analysis of congruence vs analysis of algorithmic effects

In a classificatory context, Gower (1983) distinguished between two basic types of comparisons: (1) comparing classifications obtained by the same algorithm and based on the same set of objects but on different character sets, and (2) comparing classifications that differ only in the clustering criteria used. Rohlf & Sokal (1981) referred to the first case as analysis of congruence, which they considered especially important in assessing the robustness of classifications and in tests of the nonspecificity hypothesis (Sneath & Sokal 1973). Of course, this distinction can readily be extended to all other types of OUCs.

Typical examples for the analysis of congruence in vegetation science are found in Feoli *et al.* (1984b) and Podani (1985). They evaluated classifications and ordinations derived from different character sets (e.g., floristic, chorological, pedological, and biological) for the same set of phytosociological relevés. Del Moral & Watson (1978) compared ordinations based on data obtained from tree and herb layer separately. Avena *et al.* (1981) were concerned with classifications based on binary and cover data. Such analyses of 'syntaxonomic congruence' or 'predictivity' investigate problems that are analogous to the nonspecificity problems of taxonomy. Feoli (1983) and Feoli, Orlóci & Scimone (1984b) gave a general treatment of predictivity and presented more examples.

The other alternative, the analysis of algorithmic effects on OUCs needs no comments here; this is a common practice in methodological studies (e.g., Fish 1976; Booth 1978; Robertson 1979; Gauch 1980; Gauch & Whittaker 1981; Kenkel & Orlóci 1986; Kenkel & Booth 1987; Podani 1989b) and in comparisons of results obtained by traditional and numerical methods (e.g., Stanek 1973; Mucina 1982).

Choice 6: elementary vs complex comparisons

A comparison is termed elementary when the OUCs in question have been obtained so that their differences can be attributed only to a single factor. Since this is the assumed strategy in most cases, one must clearly specify at the outset what the objectives of the study are. For instance, if the effect of clustering algorithm is to be evaluated, the other criteria of clustering (data type, resemblance functions, etc.) must be kept constant. Otherwise, the goals of the study will be difficult to achieve owing to the confounding effects of factors on OUCs (cf. Kenkel & Orlóci 1986).

In special situations, two or more underlying influential factors are simultaneously changed by design. The assessment of OUCs thus obtained implies complex (two- or multifactorial) comparisons, which may reveal the relative impact of the factors on the results. In a study of secondary succession, Virágh (1987) examined the combined effect of herbicide treatment, data type, resemblance coefficient and clustering method upon the classification of permanent plots. Examples illustrating the simultaneous effect of plot size and data type on resemblance structures, classifications and ordinations are given below in this paper.

Choice 7: univariate vs multivariate comparisons

Comparisons are usually based on a single property of OUCs; this is surprising because the previous stages of the analysis are multivariate in nature. A comparison is univariate even though $n(n-1)/2$ paired elements are involved, say, in the

comparison of two dendrograms according to their corresponding topological (or cladistic) difference matrices (in which each value corresponds to the number of internal nodes separating two objects, Phipps 1971). Also, cophenetic correlation (Sokal & Rohlf 1962) reflects dendrogram similarity only in terms of differences in hierarchical levels. The majority of methods for comparing ordinations, resemblance matrices and partitions also consider only a single aspect of the results.

Podani & Dickinson (1984) pointed out that the above strategy is less reliable if complicated OUCs, such as dendrograms, are involved in the comparison. They suggested an analytical method which incorporates as many aspects of dendrogram description as possible. The distance coefficient proposed utilizes 5 different descriptors simultaneously, each receiving equal weight, so that the comparison becomes really multivariate. It is noted that the comparisons of other types of OUCs may also be made multivariate in similar manner.

The measurement of resemblance and consensus

The methodology of comparing resemblance matrices, classifications and ordinations is a very diverse subject. An obvious reason was mentioned in the Introduction: the necessity for such work arises in a great variety of research fields; in fact, the problem is potentially present wherever multivariate techniques are utilized. The relevant information is scattered over an immense literature, and contributions to this topic substantially differ in terminology and mathematical abstraction. Therefore, instead of presenting an exhaustive review on current methodology, it seems more practical to mention major approaches only and to collect ample references to facilitate further reading.

Resemblance matrices

The multivariate analysis of vegetational entities rarely stops at the calculation of resemblance matrices; the objective is usually to obtain classi-

fications and ordinations. Nevertheless, the comparison of resemblance matrices should not be neglected since they are the only carriers of undistorted information on the resemblance structure of entities. Furthermore, techniques of matrix comparison may often be used for the evaluation of other OUCs without much modification.

There are two major approaches to the comparison of matrices. In the first case the two matrices are compared element by element using the product-moment correlation (Sokal & Rohlf 1962), a stress function (Kruskal 1964), or rank correlation (Hubert 1978). The matrix congruence test suggested by Mantel (1967) allows for a significance test of these coefficients. A more recent treatment of this topic is presented in Hubert (1983). It is always assumed that all values are equally important in the matrices, but this may not be appropriate. In the procedures suggested by Lefkovitch (1984; 1985b) dissimilarity matrices are replaced by relative neighbor graphs so that large dissimilarities are discarded from the comparison. Counting the number of edges shared by both graphs offers a similarity measure which may be tested for significance.

Ordinations

The applicability of various comparison techniques depends primarily on the number of axes to be considered. In the simplest case only one dimension is retained for each ordination and the comparison involves the measurement of correspondence between two orderings. Product-moment correlation (e.g., Gauch & Whittaker 1972; Robertson 1978), rank correlation (e.g., del Moral & Watson 1978; Fasham 1977; Feoli, Orlóci & Scimone 1984b; Robertson 1978; Minchin 1987) or simple graphical illustration of object displacement (Kessell & Whittaker 1976; Robertson 1978) may be used for this purpose. Another possibility is to calculate the mean percentage displacement of object position in one ordination relative to the other. This method was originally suggested to measure ordination efficiency in revealing one-dimensional gradients (Gauch & Whittaker 1972; Kessell & Whittaker 1976).

In two or more dimensions the above coefficients do not work, although there are attempts, for example, to average rank correlations for several axes (Feoli *et al.* 1984b), and not entirely without success. However, it is more elegant to calculate the distance between each pair of objects in every ordination and then measure ordination dissimilarity by matrix correlation techniques. Clearly, if the scaling method completely preserves the resemblance structure of objects (as PCA does), comparisons based on all dimensions and on the original resemblance matrices will provide identical results. An application of matrix correlation to compare ordination configurations was given by Gauch *et al.* (1981), who pointed to some deficiencies of this method and stressed that subjective assessment should complement the calculations.

No doubt, the dissimilarity of two ordinations may be best measured using Procrustes analysis (Schönemann & Carroll 1970; see also Gower 1971). Each of the two configurations being compared is rescaled to unit sum of squares and the least squares fit of the ordinations is found by central dilation, reflection and rigid motion of points. The error sum of squares may then be used as a dissimilarity measure (application in Fasham 1977; Kenkel & Orlóci 1986; Minchin 1987). In multiple Procrustes analysis (cf. Digby & Gower 1986) k ordinations are compared in every pair by performing $k(k-1)/2$ pairwise rotations. The resulting statistics are assembled into a symmetric matrix which can be subjected to an ordination of ordinations (see example in Digby & Kempton 1987).

There are two alternative approaches to the construction of consensus ordinations. Generalized Procrustes analysis (Gower 1975) is an iterative procedure to find a best fit for k ordinations based on raw coordinates of points. When the best fit is obtained, the average coordinates for each point will provide the consensus configuration. The residual sum of squares may be partitioned among ordinations (to show the relative departure of ordinations from the consensus) or points (to indicate objects whose position causes the greatest discrepancy among the ordinations).

If dissimilarity matrices rather than coordinates are available, individual differences scaling (INDSCAL, Carroll & Chang 1970) may be used to generate a consensus (in the 'stimulus space') and an ordination of matrices (in the 'weights space'). Kenkel & Booth (1987) provided an example of application to evaluate the performance of various binary indices in multidimensional scaling.

Partitions

Nonhierarchical classifications of objects play a central role in vegetation studies. Even though most clustering strategies used in practice generate hierarchical classifications, the investigators are in fact most interested in a partition obtained at a certain 'optimum' level in the hierarchy with an objective, for example, to create vegetation maps. Therefore, the comparison of partitions appears of central importance in quantitative vegetation surveys.

The calculation of resemblance between two partitions may start from a cross-classification table, whose entries signify numbers of objects that were clustered together in both classifications being compared. Various contingency coefficients may be applied to this table, e.g., chi-square statistics (see Anderberg 1973), information theoretical functions (Feoli *et al.* 1984a) or measures of predictivity (Goodman & Kruskal 1954). Applications of these functions to vegetation studies include Grigal & Goldstein (1971), Dale & Clifford (1976), Robertson (1979), Mucina (1982), Feoli (1983) and André (1988). Matrix comparisons may also be used to measure partition agreement. Each partition of n objects is described by an $n \times n$ incidence matrix C , in which c_{ij} is 1 if objects i and j belong to the same cluster, and 0 otherwise. Any appropriate binary resemblance coefficient may then be used to compare two such matrices element by element. For example, the simple matching coefficient was first used by Rand (1971) to evaluate partitions; while Sørensen's index was adopted by Gauch (1980) and Gauch & Whittaker (1981). It is worth

mentioning that matrix comparison measures may also be expressed in terms of the cross-classification table (Rohlf 1974), but the reverse is not true: contingency coefficients cannot be written in any other form. The third approach to the comparison of partitions involves calculating the number of elementary operations needed to transform one partition into the other ('minimum length sequence metrics', Day 1981). The simplest of such measures is the minimum number of objects which must be relocated from one group to another in one partition to reach identity to the target partition (an application is in Fish 1976). This number may be easily transformed into a measure of partition agreement (Gordon 1980). Day's (1981) sigma-metric seems to combine the three different approaches discussed so far. Podani (1986) presented a detailed treatment of this topic, including the problem of normalizing coefficients and significance tests of partition agreement, with special reference to vegetation studies. Further information is provided by Arabie & Boorman (1973), Day (1981), Gordon (1981), and Hubert & Arabie (1985).

In the construction of consensus classifications, the concept of *strict* consensus (Sokal & Rohlf 1981) is fundamental. In strict consensus partitions (or 'strong patterns', Diday & Simon 1976), any two objects are in the same class if and only if they were classified together in all competing partitions. For two classifications, the consensus groups are represented by the cells of a two-way cross-classification table; consensus classes for three partitions are summarized by a three-way table, and so on. Often, there are numerous strict consensus classes, some containing only a single object (see e.g., Podani 1989a). Thus, the strict consensus partition may be too fine from a practical viewpoint and a compromise is in order to reduce the number of classes. The majority rule (cf. Day 1988) is implicitly used in the hierarchical approaches to consensus generation: a nested system of consensus partitions is obtained by complete linkage clustering (Diday & Simon 1976) or a global optimization strategy (Podani 1989a). The structure of the consensus hierarchy is indicative of the overall agreement of

input partitions at several numbers of consensus classes. Whereas these methods are computationally efficient, the results obtained are only approximations to the optimal solution. The application of majority rule may lead to the *median* consensus partition for which the sum of distances from the input partitions is minimal. However, its direct computation for large problems is a difficult and partly unsolved problem (cf. Barthélemy & Monjardet 1988; Day 1988). A formal and exhaustive discussion of the topic is found in a special issue (vol. 3, No. 2, 1986) of the *Journal of Classification*. Although consensus partitions have attracted only little interest in vegetation studies (e.g., Pritchard & Anderson 1971; Podani 1989a), their potential usefulness for the evaluation of, say, vegetation maps is obvious.

Dendrograms

As pointed out above, in vegetation surveys the fine details of hierarchical classifications are not equally interesting. Yet, for small numbers of objects the comparison of complete dendrograms may still prove useful. As the examples will demonstrate, dendrogram comparisons may contribute much to the understanding of background effects influencing our results. Furthermore, the assessment of the performance of hierarchical clustering methods is impossible without considering all levels in the hierarchy.

Early approaches to the problem utilize matrix comparison techniques. Each dendrogram is described by an $n \times n$ matrix D , in which d_{ij} may reflect different properties of dendrograms in terms of pairwise relationships (cophenetic level, Sokal & Rohlf 1962; cladistic – or 'topological' – difference, Phipps 1971; and three other descriptors suggested by Podani & Dickinson 1984). Product-moment correlation (Sokal & Rohlf 1962), rank correlation and stress functions (Jackson 1969; Cunningham & Ogilvie 1972), Euclidean distance (Phipps 1971), or the INDSCAL method (Rohlf 1974) are then applied to compare two descriptor matrices. Podani & Dickinson (1984) suggested incorporation of all the five descriptors in the same distance measure.

Of course, dendrogram comparison is not necessarily based on the description of the relative position of object pairs in the hierarchy. Dobson (1975) proposed to count the number of object triplets in one dendrogram for which the ultrametric inequalities are different in the other dendrogram. This number divided by the possible number of object triplets provides a dissimilarity measure. The underlying principle of other methods involves the deletion of interior edges from the dendrogram to yield subtrees and partitions, which are subsequently compared. Dobson (1975) suggested to divide the number of common subtrees by the number of different subtrees in the two dendrograms, but this ratio is very sensitive to small differences in the hierarchies. Farris (1973) introduced the cluster distortion method which operates by counting the number of fragments in one dendrogram into which a subtree of the other dendrogram is demolished. An average for all subtrees will give an asymmetric measure. Dale & Moon (1988) described a method for testing the departure of dendrogram structure from random expectation. They suggest that this test can be used for comparisons, but no details are presented. Robinson & Foulds (1979, 1981) pointed out that some elementary operations which transform one phylogenetic tree into another may be used to define distances between trees. They showed that the minimum number of such operations required to obtain one tree from the other is the number of edges whose removal results in a partition which is present in only one of the two trees (mismatched edges). Although the elementary operations do not apply to dendrograms, the concept of mismatched edges seems useful for dendrogram comparisons. The number of mismatched edges divided by the total number of edges (i.e., $2n - 4$) provides a dissimilarity measure (edge matching coefficient). Other indices may be constructed by assigning weights to the edges according to the hierarchical levels (absolute edge difference and mismatched edge difference).

Other methods reduce the problem to the comparison of partitions. Fowlkes & Mallows (1980) proposed a method which operates with the com-

parison of two dendrograms at all possible levels and portrays the similarities in a profile diagram. The efficiency of this procedure is doubtful, however, and the results are dependent on the properties (e.g., actual ranges) of the resemblance coefficient used.

There are several approaches to the construction of consensus OUCs from dendrograms. Strict consensus trees (Sokal & Rohlf 1981) contain only those clusters that are found in all the original dendrograms while no values are assigned to hierarchical levels. Consensus dendrograms are generated from matrices of cophenetic levels (ultrametrics) such that each pair of objects joins at a dissimilarity level corresponding to the maximum of levels for that pair in the dendrograms (see Faith 1988). Lefkovitch (1985a) adopted a different approach; he uses principal coordinates analysis of these matrices and the resulting point configurations are fitted by Procrustes analysis. Then, signs of the consensus coordinates are used to define the consensus hierarchy (the method also applies to partitions). The 'pruning' algorithm of Gordon (1980) removes branches from the dendrograms to obtain a reduced tree which does not conflict with the original dendrograms. Then, reattachment of the removed branches into some 'average' positions gives a 'grafted' tree showing which objects are responsible for agreements and disagreements among the classifications.

Consensus indices (Rohlf 1982; Faith 1988) may be used to derive dissimilarities from consensus dendrograms or trees. The distributional properties of some consensus indices are examined by Shao & Rohlf (1983). The intermediate measures of Faith & Belbin (1986) are derived from a consensus classification and metric distances between the dendrograms.

Examples

All cases to be presented here will illustrate the *exploratory* approach to the *intra-type* comparison of OUCs for evaluating *congruence* of results. Cases 1–3 involve *complex, overall* comparisons, in case 4 *elementary* comparisons are made using

a standard OUC as a *reference* basis. In all these cases *resemblance* of OUCs is computed, whereas case 5 exemplifies *consensus* generation for ordinations. Partitions of vegetation data are not analyzed here, such studies are included in Podani (1985, 1986, 1989a) using the same data set described below. All calculations were performed by the programs from the SYN-TAX III package (Podani 1988).

Field data

In 1977, 8 sets of relevés were made in the grassland communities on dolomite in the Sashegy Nature Reserve, Hungary. This area of ca. 30 ha lies within the city limits of Budapest at ca. 350 m above sea level. Zólyomi (1958) presented a detailed description of its vegetation using traditional phytosociological methods. 80 quadrats were placed in the undisturbed parts of the study area. Each quadrat represents a nested system of 8 square plots increasing in size with a common corner. The size of the smallest plots was 0.5×0.5 m, and they were enlarged by 0.5 m increments on both sides up to the size of 4×4 m. The corners of each plot were permanently marked with large steel pegs so as to make possible to check the site at different times from early spring to autumn. Percentage cover estimates for vascular plants were recorded. The number of species at increasing quadrat sizes was respectively 79, 101, 108, 110, 112, 118, 121, and 123. These values, plotted against sampling unit size, clearly indicate the existence of logarithmic species/area relationship which has previously been reported in many studies.

Case 1: The effect of data type and plot size on dendrograms

Two series of classifications of the 80 quadrats, each series consisting of 8 dendrograms based on different plot sizes, were obtained by the incremental sum of squares method (SSA, Orlóci 1967). In the first series presence/absence scores,

in the other series cover data standardized by the range of each species were used. A distance matrix for the 16 dendrograms was calculated utilizing cophenetic differences, cladistic differences and cluster membership divergences as descriptors of dendrogram structure (see Podani & Dickinson 1984, for details of the methods). This matrix was subjected to SSA and principal coordinates analysis (PCoA) to evaluate the simultaneous influence of plot size and data type upon dendrograms.

In addition, the 45 most common species, with frequency higher than 6, were classified based on presence/absence and cover to yield other 2 series of dendrograms. The data were standardized by the range of cover values within each quadrat. The 16 species classifications were evaluated in the same manner as the dendrograms of quadrats.

Most classifications of quadrats on presence/absence basis are quite similar to one another (Fig. 1a). Only the classification pertaining to plot size 3.5×3.5 m (No. 7 in the figure) falls relatively far apart. Within this series, the trend of increasing plot sizes is obscured suggesting that random variation is at least as important as sampling. The dendrograms based on cover data exhibit more considerable differences, as they are positioned around the binary series and their arrangement seems to reflect plot size increases pretty well. Groups derived by SSA correspond to small (1–3), intermediate (4–6) and large (7–8) plot sizes, and are located along an arch in the ordination.

Whereas quadrat classifications proved more stable in the presence/absence situation, species classifications exhibit the opposite behaviour (Fig. 1b). The more stable classifications resulted when cover data were used, at the two-cluster level the groupings did not change at all (not illustrated). In the binary case, however, increases in plot size induced a continuous rearrangement of the hierarchy. The linearity is not distorted by the arch effect, which may be due to the presence of the compact cluster at the other end of axis 1. It is also striking that the changes are more substantial at small sizes, that is, the classifications approximate a fairly stable situation as the plots

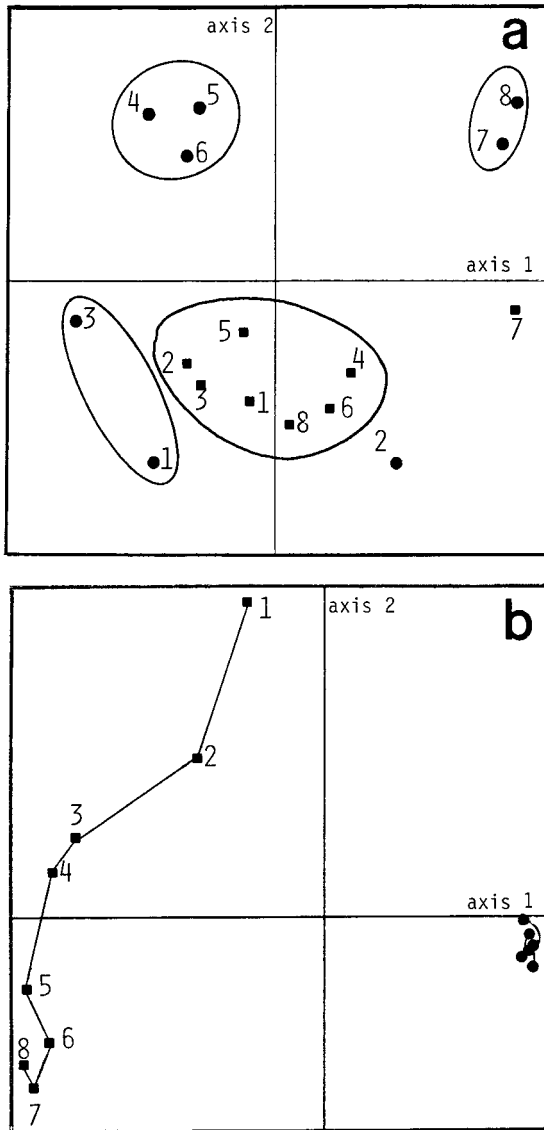


Fig. 1. Principal coordinates analysis of the combined effect of plot size and data type on SSA classifications of 80 quadrats (a) and 45 species (b). Points identify dendrograms based on data according to the symbols (■ binary data, ● percentage cover standardized by range) and numbers (plot size increases from 1 to 8). Loops indicate groups obtained by SSA of dendrograms.

are enlarged. This is because the sampling units become more and more 'saturated' with the most common species that potentially occur at the given site.

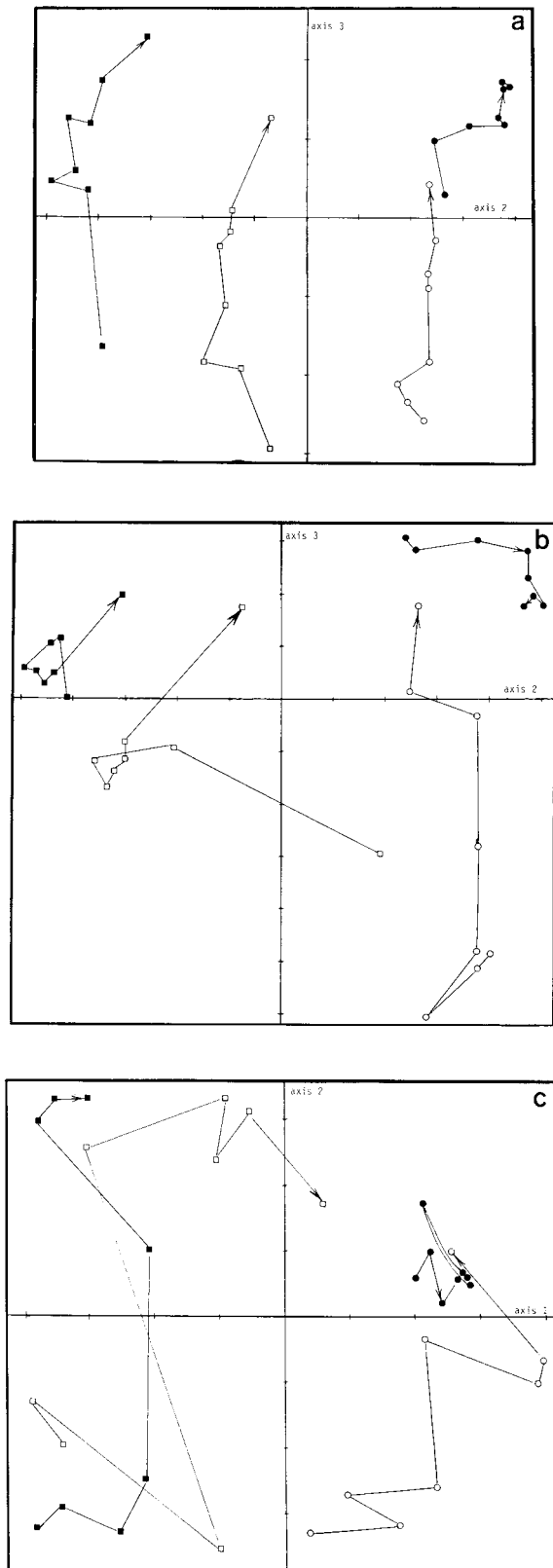
Although only 45 species out of the 123 were included in the classification, such a discrepancy between the results is unexpected. A possible

explanation is that it is the continuous change of species classifications which ensures relative stability of quadrat classifications for the binary data. The high stability of species classifications using cover data may be partly attributed to the standardization procedure (i.e., division by the maximum score for each quadrat).

Case 2: The effect of data type and plot size on resemblance structures, ordinations and classifications

In this example the 8 plot sizes are combined with 4 data types, yielding a total of 32 OUCs. To simplify the analysis, 19 nested plots containing a total of 104 species were selected from the full sample. 4 data types were considered: percentage cover, transformed cover data obtained by Clymo's function ($x' = (1 - \exp(-cx)) / (1 - \exp(-c))$, see van der Maarel 1979) with $c = 3$ and $c = 15$, and binary scores. These operations represent a data transformation process (Podani 1984). Cover and binary data are extremes of the series, the case with $c = 3$ is closer to the cover data whereas $c = 15$ yields an approximation to the presence/absence data. Euclidean distance between the 19 quadrats was computed in each combination of plot size and data type. Multivariate analysis of quadrats was performed by SSA and PCoA.

The distance matrices were compared by the product-moment correlation coefficient, and the resulting correlation matrix was analyzed by PCA. The first three components explained 84%, 9% and 2.3% of the total variance. Component 1 is strongly unipolar with scores ranging from 0.8 to 0.96. Therefore, a scattergram is given only for the second and third components (Fig. 2a). This ordination shows the existence of a two-dimensional continuum resulting from the joint effect of two factors. Axis 2 can be identified as the data type factor, axis 3 corresponds to the plot size effects. The ratio of the associated eigenvalues suggests that data type has about 4 times more influence on the resemblance structure of quadrats than plot size. The change of re-



semblance structure due to plot size is the least substantial for the presence/absence data.

The 32 ordinations of quadrats were analyzed by PCoA based on the distances of points in the ordination space. Since in most phytosociological surveys only 2 or 3 dimensions are evaluated and illustrated, the comparisons were also restricted to 2 and 3 dimensions. For 2 dimensions, the percentages of first 3 eigenvalues were 63%, 14%, and 7% respectively. Similarly as in the ordination of distance matrices, the first component was unipolar, explaining a large amount of shared variation in the OUCs. However, the next 2 axes depict a meaningful configuration (Fig. 2b). The components can also be identified as data type and plot size factors, but the trends are less obvious. Another difference from the ordination in Fig. 2a is that the extreme data types yield relatively stable ordinations, whereas transformed data generate more divergent results. As expected, when 3 dimensions are considered the ordination of ordinations (not illustrated) is much closer to the ordination of distance matrices.

The dendrograms were analyzed utilizing 5 dendrogram descriptors and PCoA. The first component is not unipolar, and explained 21% of the variance, the second 2 axes account for 16% and 10%, respectively. The configuration is the most informative in the first 2 dimensions (Fig. 2c). The picture is less clear than in the other 2 ordinations, although the data type and plot size gradients are still recognizable with the intermediate data types mixed up with the extremes. The reason for the distortion is two-fold. The description of dendrograms is still incomplete with 5 descriptors, so that ordination efficiency could be improved by incorporating further descriptors (to be developed in the future). Despite this incompleteness, the ordination demonstrates

Fig. 2. Analysis of the combined effect of 4 data types and 8 plot sizes on OUCs. Points identify resemblance matrices (a), PCoA ordinations (b), and SSA dendrograms (c) for 19 quadrats based on data according to the symbols: ● binary data, ○ Clymo's transformation with $c = 15$, □ Clymo's transformation with $c = 3$, and ■ percentage cover. Arrows indicate increase in plot size.

that resemblance structures are distorted by the clustering algorithm. The relative stability of binary data is also supported by this result; for all other data types the effect of plot size was incomparably more substantial. This confirms the conclusions drawn in Case 1.

Case 3: The effect of data standardization and taxonomic level on dendrograms

The applicability of higher taxonomic ranks to phytosociological classifications was examined using 80 4 × 4 m quadrats. Raw and range-standardized percentage cover were combined with 4 different levels (species, genus, family, and order). The successive fusion of taxa followed Engler's (1964) system. The number of taxa at the 4 levels was 123, 96, 32, and 17, respectively. The quadrats were classified by SSA and the resulting 8 dendrograms were ordinated by PCoA using 5 descriptors.

The lowest 2 taxonomic categories produced very similar classifications, as well as the families and orders (Fig. 3). Standardization by range did not affect these relationships; in both cases the combination of genera into families modified the OUCs most significantly. This step represents the most drastic reduction in the number of variables (from 96 to 32); other changes are smaller. Note that contrast between dendrograms of standardized data is explained by axis 1, whereas the dendrograms of raw data are separated on axis 2. A conclusion is that reduction of taxonomic ranks is most critical for the standardized case.

Case 4: The effect of data standardization and number of species on dendrograms

19 4 × 4 m quadrats were selected from the sample. Two different data sets were used; raw cover and cover standardized by range for each species. The total error sum of squares was calculated for both types. In each data set the species were ranked according to their contribution to the total sum of squares such that correlated species

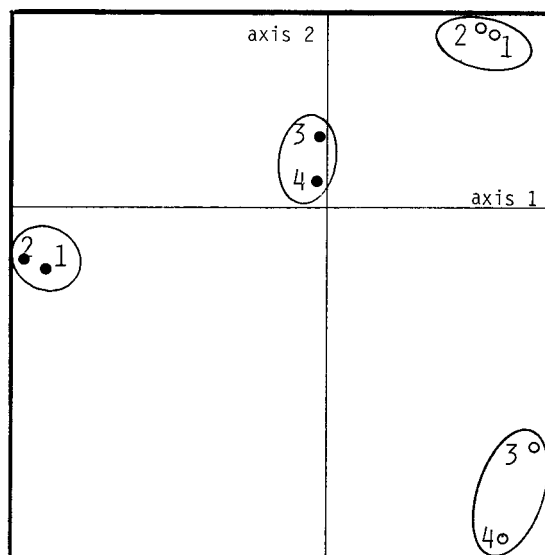


Fig. 3. PCoA analysis of the combined effect of taxonomic level and data standardization on dendrograms. Each point represents an SSA classification of 80 4 × 4 m quadrats based on different level (1. species, 2. genus, 3. family, and 4. order) and data type (o raw cover, • standardized cover).

were not eliminated. This ranking technique is compatible with the SSA method subsequently used for the clustering of quadrats. For raw data the classifications were based on the first 104 (100%), 75 (99.8%), 50 (99.4%), 25 (98%), 10 (91%), and 5 (79.8%) species in the rank order (values in brackets are the pooled sum of squares as a percentage of total). In the case of standardized data, 8 levels of data reduction were considered. The numbers of species retained and associated percentages are: 104 (100%), 93 (95%), 83 (90%), 63 (80%), 48 (70%), 36 (60%), 26 (50%), and 19 (40%).

As a reference for comparison, the classifications based on 104 species were used. 7 coefficients of dendrogram dissimilarity were calculated: Euclidean distance between cophenetic difference matrices (ECP), Euclidean distance between cladistic difference matrices (ETD), Euclidean distance using cluster membership divergence (ECM), ultrametric dissimilarity (UD), edge matching coefficient (EM), absolute edge difference (ED), and mismatched edge difference (MED).

The relationships between the number of species used and the dissimilarity of dendrograms to the reference OUC are displayed by the profile diagrams in Fig. 4. The values are rescaled to the actual maximum in each method, and the points are connected to visualize trends but the different functions become comparable.

The two sets of curves demonstrate that the effect of species deletion varies with the data type

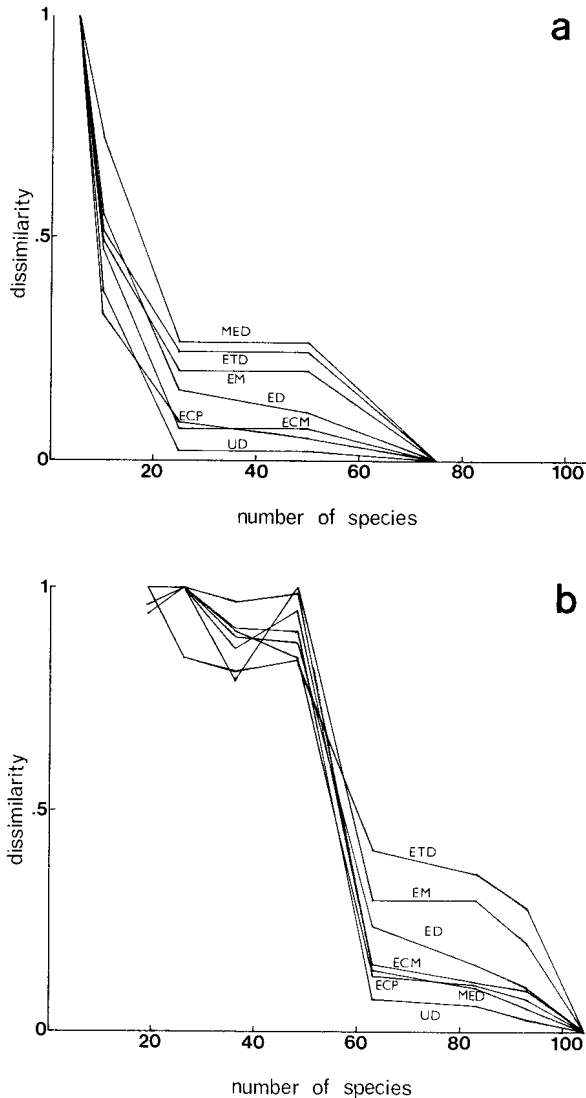


Fig. 4. The effect of number species on SSA classifications of 19 4×4 m quadrats. Profile diagrams indicate the dissimilarities to the classification based on all species as the number of species is reduced from a sum of squares rank order. a) Raw data. b) Standardized data. See text for explanation of abbreviations of functions.

used. For unstandardized data (Fig. 4a) at least the 29 lowest-ranked species may be deleted without changing the dendrograms. Deletion of further 25 species causes minor changes in the hierarchy. The tree topology is preserved upon the elimination of 25 more species, only the levels are modified as the increase of ED and ECP shows. The use of as few as 10 species and, in particular, 5 species leads to relatively drastic jumps in the dissimilarities, but on the absolute scale these values are still low. Visual inspection of the dendrograms suggests that even 5 species produced an acceptable classification. This is because these 5 species themselves explain a high percentage of variation.

Data standardization tends to balance species importance as the differences between Figs 4a and 4b demonstrate. In fact, the actual dissimilarities are higher for 48 species in the standardized case than for 5 species in the unstandardized case. The flattening of curves shows that when the number of species falls under 48, the dissimilarities are no longer increased, although the theoretical maximum (where known) is not reached (e.g., EM = 0.588). Unfortunately, it was not possible to follow this process further, because the classification based on 13 species was not binary in many parts of the dendrogram. An obvious conclusion from this study is that percentages rather than numbers of species determine ranking efficiency.

The diagrams also allow for the comparison of the 7 coefficients. UD, ECM and ECP are consistently lower than ETD, EM and ED; the differences on the relativized scale may be very high. Only MED was found as an exception to this tendency. It may be concluded that whereas ETD, EM and ED are too conservative when minor changes occur, ECM, ECP and UD appear more sensitive to major changes in the hierarchy.

Case 5: The effect of plot size on ordinations

The nested quadrats were analyzed by PCoA from Euclidean distance matrices to obtain a series of 8 ordinations. Then, generalized Pro-

crustes analysis was used to create a consensus configuration. The consensus OUC is not shown here, more interesting is the partitioning of the total error sum of squares among the original ordinations as well as among the quadrats.

The departure of ordinations from the consensus is shown in Fig. 5. Quadrats of very small size (0.25 and 1 m²) produced the most divergent results; they are responsible for a total of 48% of the sum of squares. Increases in quadrat size lead to a clear, although not monotonic decrease of the percentages, reaching the minimum at 6.25 m² (only 5.2%). Further quadrat size changes result in monotonic increases of the percentages. Although it may be expected that some 'intermediate' plot size will provide the ordination closest to the consensus configuration, the differences between the two endpoints of the plot size gradient is striking. This suggests that ordination stability is not equally influenced by plot size increments; changes are most critical for small quadrats. The contribution of individual quadrats to the sum of squares is uneven. 6 quadrats are shifted considerably as plot size increases; these account for 36.3% of the sum of squares. Each of the other 74 quadrats explains less than 3%. Removal of the inconsistently positioned 6 quadrats would probably increase ordination stability.

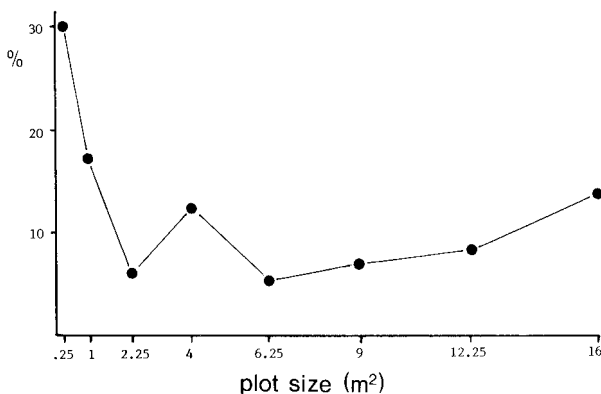


Fig. 5. Percentage contribution of PCoA ordinations of 80 quadrats, with increasing plot size, to the error sum of squares in generalized Procrustes analysis.

Concluding remarks

The review of current methodology and published comparisons of vegetation analyses reveals that applications are far behind the development of new techniques, as in taxonomy (cf. Faith 1988). Vegetation science utilizes only a narrow range of methods, most of them taken from conventional statistics (e.g., measures of correlation) or simply adopted uncritically from other fields of application (e.g., some binary indices for expressing partition agreement). The time-lag is not less striking in comparison with numerical taxonomy. Consensus methods, for example, have become standard data analytical tools in phenetic and cladistic studies, but their potential usefulness in community studies is little investigated.

Most published comparisons in vegetation science place emphasis on the evaluation of the performance of multivariate procedures. It is repeatedly stressed in this paper, however, that our field has many specific problems which deserve much more attention. Succession studies and analyses of congruence between ecological and floristical data may benefit from comparative assessment of results. Examples demonstrated the utility of comparisons in evaluating the relative impact of decisions made during community studies. Possibilities of such work are by no means exhausted, and answers to questions are not necessarily trivial as it might appear from some of the examples presented above. First of all, problems with optimal sampling designs would require an extensive study using the multiple comparison approach. Also, a survey would be most useful to compare the different cover-abundance scales and data standardization methods in classification and ordination context. In general, comparison of results is recommended in any survey where interest lies in examining changes of ordinations and classifications of vegetation data.

Acknowledgements

I am grateful to the Soros Foundation (New York – Budapest) for financial support to attend the

conference in Vienna. I thank L. Orlóci and P. Juhász-Nagy for comments on the manuscript, and several people at the conference for raising problems and questions. Computing facilities were provided by the University of Western Ontario and the Hungarian Academy of Sciences.

References

- Adams III, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.* 21: 390–397.
- Anderberg, M. R. 1973. Cluster analysis for applications. Academic Press, New York.
- André, H. M. 1988. Variable centered methods and community classification. *Coenoses* 3: 69–78.
- Arabie, P. & Boorman, S. A. 1973. Multidimensional scaling of measures of distance between partitions. *J. Math. Psychol.* 10: 148–203.
- Austin, M. P. 1976. Performance of four ordination techniques assuming different non-linear species response models. *Vegetatio* 33: 43–49.
- Avena, G., Blasi, C., Feoli, E. & Scoppola, A. 1981. Measurement of the predictive value of species lists for species cover in phytosociological samples. *Vegetatio* 45: 77–84.
- Barthélemy, J. P. & Monjardet, B. 1988. The median procedure in data analysis: new results and open problems. In: Bock, H. H. (ed.), *Classification and related methods of data analysis*. pp. 309–316. North-Holland, Amsterdam.
- Bayne, C. K., Beauchamp, J. J., Begovich, C. L. & Kane, V. E. 1980. Monte Carlo comparisons of selected clustering procedures. *Pattern Recognition* 12: 51–62.
- Booth, T. H. 1978. Numerical classification techniques applied to forest tree distribution data. I. A comparison of methods. *Australian J. Ecol.* 3: 297–306.
- Burgman, M. A. 1987. An analysis of the distribution of plants on granite outcrops in southern Western Australia using Mantel tests. *Vegetatio* 71: 79–86.
- Carroll, J. D. & Chang, J. J. 1970. Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition. *Psychometrika* 35: 283–319.
- Crovello, T. J. 1981. Quantitative biogeography: an overview. *Taxon* 30: 563–575.
- Cunningham, K. M. & Ogilvie, J. C. 1972. Evaluation of hierarchical grouping techniques. A preliminary study. *Computer J.* 15: 209–213.
- Dale, M. B. & Clifford, H. T. 1976. On the effectiveness of higher taxonomic ranks for vegetation analysis. *Australian J. Ecol.* 1: 37–62.
- Dale, M. R. T. & Moon, J. W. 1988. Statistical tests on two characteristics of the shapes of cluster diagrams. *J. Classification* 5: 21–38.
- Day, W. H. E. 1981. The complexity of computing metric distances between partitions. *Math. Soc. Sci.* 1: 269–287.
- Day, W. H. E. 1983. Distributions of distances between pairs of classifications. In: Felsenstein, J. (ed.), *Numerical taxonomy*, pp. 127–131. Springer, Berlin.
- Day, W. H. E. 1988. Consensus methods as tools for data analysis. In: Bock, H. H. (ed.), *Classification and related methods of data analysis*, pp. 317–324. North-Holland, Amsterdam.
- del Moral, R. & Watson, A. F. 1978. Gradient structure of forest vegetation in the central Washington Cascades. *Vegetatio* 32: 29–48.
- Dickinson, T. A., Knowles, P. & Parker, W. H. 1988. Data set congruence in northern Ontario tamarack (*Larix laricina*, *Pinaceae*). *Syst. Bot.* 13: 442–455.
- Diday, E. & Simon, J. C. 1976. Clustering analysis. In: Fu, K. S. (ed.), *Digital pattern recognition*, pp. 47–94. Springer, New York.
- Dietz, E. J. 1983. Permutation tests for association between two distance matrices. *Syst. Zool.* 32: 21–26.
- Digby, P. G. N. & Gower, J. C. 1986. Ordination and classification. Univ. Montreal Press, Montreal.
- Digby, P. G. N. & Kempton, R. A. 1987. Multivariate analysis of ecological communities. Chapman & Hall, London.
- Dobson, A. J. 1975. Comparing the shapes of trees. In: Dold, A. & Eckmann, B. (eds), *Lecture notes in mathematics* 452: 95–100. Springer, Berlin.
- Dubes, R. & Jain, A. K. 1976. Clustering techniques: the user's dilemma. *Pattern Recognition* 8: 247–260.
- Duncan, T. & Estabrook, G. F. 1976. An operational method for evaluating classifications. *Syst. Bot.* 1: 373–382.
- Duncan, T., Phillips, R. B. & Wagner, W. H. 1980. A comparison of branching diagrams derived by various phenetic and cladistic methods. *Syst. Bot.* 5: 264–293.
- Engler, A. 1964. *Syllabus der Pflanzenfamilien*. Gebrüder Borntraeger, Berlin.
- Faith, D. P. 1988. Consensus applications in the biological sciences. In: Bock, H. H. (ed.), *Classification and related methods of data analysis*, pp. 325–332. North-Holland, Amsterdam.
- Faith, D. P. & Belbin, L. 1986. Comparison of classifications using measures intermediate between metric dissimilarity and consensus similarity. *J. Classification* 3: 257–280.
- Farris, J. S. 1973. On comparing the shapes of taxonomic trees. *Syst. Zool.* 22: 50–54.
- Fasham, M. J. R. 1977. A comparison of nonmetric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines, and coenoplanes. *Ecology* 58: 551–561.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
- Feoli, E. 1983. Predictive use of classification and ordination methods in plant community ecology. A summary with examples. In: Ferrari, G., Gentile, S., Pignatti, S. & Poli Marchese, E. (eds), *Le comunità vegetali come indicatori ambientali*, pp. 83–108. Bologna.
- Feoli, E., Lagonegro, M. & Orlóci, L. 1984a. Information analysis of vegetation data. Junk, The Hague.

- Feoli, E., Orlóci, L. & Scimone, M. 1984b. Comparison of ordinations of some Apennine forest communities based on different characters and methods. *Studia Geobot.* 4: 137–145.
- Ferraris, J. 1983. Comparison of classifications with the data from which they are derived. In: Felsenstein, J. (ed.), *Numerical taxonomy*, pp. 174–178. Springer, Berlin.
- Fish, E. B. 1976. Comparisons of phytosociological methods of classification on a desert grassland site. *Agro-Ecosystems* 2: 173–194.
- Fowlkes, E. B. & Mallows, C. L. 1980. A method for comparing two hierarchical clusterings. Manuscript, Bell Laboratories, Murray Hill, N. J., USA. Abstract in: *Classification Soc. Bull.* 4: 54.
- Gauch, H. G. 1982. Multivariate analysis of community data. Cambridge Univ. Press, Cambridge, Mass., USA.
- Gauch, H. G. 1980. Rapid initial clustering of large data sets. *Vegetatio* 42: 103–111.
- Gauch, H. G. & Whittaker, R. H. 1972. Comparison of ordination techniques. *Ecology* 53: 868–875.
- Gauch, H. G. & Whittaker, R. H. 1981. Hierarchical classification of community data. *J. Ecol.* 69: 537–557.
- Gauch, H. G., Whittaker, R. H. & Wentworth, T. R. 1977. A comparative study of reciprocal averaging and other ordination techniques. *J. Ecol.* 65: 157–174.
- Goodman, L. A. & Kruskal, W. H. 1954. Measures of association for cross-classifications. *J. Amer. Stat. Assoc.* 49: 732–764.
- Gordon, A. D. 1980. On the assessment and comparison of classifications. In: Tomassone, R. (ed.), *Analyse de données et informatique*, pp. 161–171. I.R.I.A., Le Chesnay.
- Gordon, A. D. 1981. *Classification*. Chapman & Hall, London.
- Gower, J. C. 1971. Statistical methods of comparing different multivariate analyses of the same data. In: Hodson, F. R., Kendall, D. G. & Tautu, P. (eds), *Mathematics in the archaeological sciences*, pp. 138–149. Edinburgh Univ. Press, Edinburgh.
- Gower, J. C. 1975. Generalized Procrustes analysis. *Psychometrika* 40: 33–51.
- Gower, J. C. 1983. Comparing classifications. In: Felsenstein, J. (ed.), *Numerical taxonomy*, pp. 137–155. Springer, Berlin.
- Greig-Smith, P. 1983. *Quantitative plant ecology*. 3rd ed. Blackwell, Oxford.
- Grigal, D. F. & Goldstein, R. A. 1971. An integrated ordination-classification analysis of an intensively sampled oak-hickory forest. *J. Ecol.* 59: 481–492.
- Hubert, L. J. 1978. Generalized proximity function comparisons. *Br. J. Math. Statist. Psychol.* 31: 179–192.
- Hubert, L. J. 1983. Inference procedures for the evaluation and comparison of proximity matrices. In: Felsenstein, J. (ed.), *Numerical taxonomy*, pp. 209–228. Springer, Berlin.
- Hubert, L. & Arabie, P. 1985. Comparing partitions. *J. Classification* 2: 193–218.
- Jackson, D. N. 1969. Comparison of classifications. In: Cole, A. J. (ed.), *Numerical taxonomy*, pp. 91–113. Academic Press, New York.
- Jeglum, J. K., Wehrhahn, C. F. & Swan, J. M. A. 1971. Comparisons of environmental ordinations with principal component vegetational ordinations for sets of data having different degrees of complexity. *Can. J. Forest Res.* 1: 99–112.
- Kenkel, N. C. & Booth, T. 1987. A comparison of presence-absence resemblance coefficients for use in biogeographical studies. *Coenoses* 2: 25–30.
- Kenkel, N. C. & Orlóci, L. 1986. Applying metric and non-metric multidimensional scaling to ecological studies: some new results. *Ecology* 67: 919–928.
- Kershaw, K. A. & Looney, J. H. H. 1985. *Quantitative and dynamic plant ecology*. 3rd ed. E. Arnold, London.
- Kessell, S. R. & Whittaker, R. H. 1976. Comparison of three ordination techniques. *Vegetatio* 32: 21–29.
- Kruskal, J. B. 1964. Nonmetric multidimensional scaling: a numerical method. *Psychometrika* 29: 115–129.
- Lefkovich, L. P. 1984. A nonparametric method for comparing dissimilarity matrices, a general measure of biogeographic distance, and their application. *Am. Nat.* 123: 484–499.
- Lefkovich, L. P. 1985a. Euclidean consensus dendrograms and other classification structures. *Math. Biosci.* 74: 1–15.
- Lefkovich, L. P. 1985b. Further nonparametric tests for comparing dissimilarity matrices based on the relative neighborhood graph. *Math. Biosci.* 73: 71–88.
- Legendre, L. & Legendre, P. 1983. *Numerical ecology*. Elsevier, Amsterdam.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209–220.
- Milligan, G. W. & Isaac, P. D. 1980. The validation of four ultrametric clustering algorithms. *Pattern Recognition* 12: 41–50.
- Minchin, P. R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio* 69: 89–107.
- Mucina, L. 1982. Numerical classification and ordination of ruderal plant communities (*Sisymbrietalia*, *Onopordetalia*) in the western part of Slovakia. *Vegetatio* 48: 267–275.
- Orlóci, L. 1976. An agglomerative method for classification of plant communities. *J. Ecol.* 55: 193–205.
- Orlóci, L. 1978. *Multivariate analysis in vegetation research*. 2nd ed. Junk, The Hague.
- Orlóci, L., Kenkel, N. C. & Fewster, P. H. 1984. Probing simulated vegetation data for complex trends by linear and nonlinear ordination methods. *Abstr. Bot.* 8: 163–172.
- Orlóci, L. & Mukkattu, M. M. 1973. The effect of species number and type of data on the resemblance structure of a phytosociological collection. *J. Ecol.* 61: 37–46.
- Paule, L. & Gömöry, D. 1987. Comparison of different clustering methods applied to matrices of genetic distances based on allelic frequencies. *Abstr. Bot.* 11: 43–51.
- Penny, D., Foulds, L. R. & Hendy, M. D. 1982. Testing the

- theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. *Nature* 297: 197–200.
- Phipps, J. B. 1971. Dendrogram topology. *Syst. Zool.* 20: 306–308.
- Pielou, E. C. 1984. The interpretation of ecological data. Wiley, New York.
- Podani, J. 1984. Spatial processes in the analysis of vegetation: theory and review. *Acta Bot. Hung.* 30: 75–118.
- Podani, J. 1985. Syntaxonomic congruence in a small-scale vegetation survey. *Abstr. Bot.* 9: 99–128.
- Podani, J. 1986. Comparison of partitions in vegetation studies. *Abstr. Bot.* 10: 235–290.
- Podani, J. 1988. SYN-TAX III. A package of programs for data analysis in community ecology and systematics. *Coenoses* 3: 111–119.
- Podani, J. 1989a. A method for generating consensus partitions and its application to community classification. *Coenoses* 4: 1–10.
- Podani, J. 1989b. New combinatorial clustering methods. *Vegetatio* 81: 61–77.
- Podani, J. & Dickinson, T. A. 1984. Comparison of dendrograms: a multivariate approach. *Can. J. Bot.* 62: 2765–2778.
- Pritchard, N. M. & Anderson, A. J. B. 1971. Observations on the use of cluster analysis in botany with an ecological example. *J. Ecol.* 59: 727–747.
- Rand, W. M. 1971. Objective criteria for the evaluation of clustering methods. *J. Amer. Stat. Assoc.* 66: 846–850.
- Robertson, P. A. 1978. Comparisons of techniques for ordinating and classifying old-growth floodplain forests in southern Illinois. *Vegetatio* 37: 43–51.
- Robertson, P. A. 1979. Comparisons among three hierarchical classification techniques using simulated coenoplanes. *Vegetatio* 40: 175–183.
- Robinson, D. F. & Foulds, L. R. 1979. Comparison of weighted labelled trees. In: Horadam, A. F. & Wallis, W. D. (eds), *Lecture notes in mathematics* 748: 119–126. Springer, Berlin.
- Robinson, D. F. & Foulds, L. R. 1981. Comparison of phylogenetic trees. *Math. Biosci.* 53: 131–147.
- Rohlf, F. J. 1974. Methods of comparing classifications. *Annual Rev. Ecol. Syst.* 5: 101–113.
- Rohlf, F. J. 1982. Consensus indices for comparing classifications. *Math. Biosci.* 59: 131–144.
- Rohlf, F. J. & Sokal, R. R. 1981. Comparing numerical taxonomic studies. *Syst. Zool.* 30: 459–490.
- Schnell, G. D. 1970. A phenetic study of the suborder *Lari* (*Aves*). II. Phenograms, discussion and conclusions. *Syst. Zool.* 19: 264–302.
- Schönemann, P. H. & Carroll, R. M. 1970. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika* 35: 245–256.
- Shao, K. & Rohlf, F. J. 1983. Sampling distribution of consensus indices when all bifurcating trees are equally likely. In: Felsenstein, J. (ed.), *Numerical taxonomy*, pp. 132–136. Springer, Berlin.
- Sneath, P. H. A. & Sokal, R. R. 1973. *Numerical taxonomy*. 2nd ed. Freeman, San Francisco.
- Sokal, R. R. 1979. Testing statistical significance of geographic variation patterns. *Syst. Zool.* 28: 227–232.
- Sokal, R. R. & Rohlf, F. J. 1962. The comparison of dendrograms by objective methods. *Taxon* 11: 33–40.
- Sokal, R. R. & Rohlf, F. J. 1981. Taxonomic congruence in the *Leptodomorpha* re-examined. *Syst. Zool.* 30: 309–324.
- Stanek, W. 1973. A comparison of Braun-Blanquet's method with sum of squares agglomeration for vegetation classification. *Vegetatio* 27: 323–345.
- van der Maarel, E. 1979. Multivariate methods in phytosociology, with reference to the Netherlands. In: Werger, M. J. A. (ed.), *The study of vegetation*, pp. 161–225. Junk, The Hague.
- Virágh, K. 1987. The effect of herbicides on vegetation dynamics: comparison of classifications. *Abstr. Bot.* 11: 53–70.
- Zólyomi, B. 1958. The natural vegetation of Budapest and its environs. In: Pécsi, M. (ed.), *Natural history of Budapest (in Hungarian)*, pp. 508–642. Akadémiai, Budapest.