EVALUATION OF ORDINATION METHODS THROUGH SIMULATED COENOCLINES: SOME COMMENTS*

Enrico FEOLI & Laura FEOLI CHIAPELLA**

Istituto ed Orto Botanico, Università degli Studi, Via A. Valerio 30, Trieste, Italy 34100

Keywords:
Coenocline, Linearity, Ordination, Seriation

Introduction

The ordinations based on resemblance matrices (Orlóci 1978b) received special attention in plant ecology during the past decade. Many papers dealing with this topic could be mentioned. It should however be satisfactory to refer to the general survey by Dale (1975). He distinguished two main approaches in ordination (cf. also Beals 1973): One is associated with pattern recognition and pattern analysis (ordination sensu lato, Noy-Meir & Whittaker 1977), which started with Goodall (1954) and is still dominant in phytosociology (see Westhoff & van der Maarel 1978 and Orlóci 1978a for references). The other one is more strictly related to Whittaker's work (ordination sensu stricto, Noy-Meir & Whittaker 1977) and concerned with gradient analysis (Whittaker 1967).

As stated by Austin (1976a) these two approaches are complementary since the former is the basis for the latter. The fact that both are productive and not competitive has been emphasized by van der Maarel (1971), and Gauch, Chase & Whittaker (1974). However, a strong tendency exists to evaluate the ordination methods mainly on the basis of their capacity to give axes easily interpretable in terms of gradients. It has in fact been suggested to evaluate the performance of different ordination methods through simulated coenoclines (Swan 1970, Noy-Meir & Austin 1970, Austin & Noy-Meir 1971, Gauch & Whittaker 1972a, b, Kessel & Whittaker 1976, Austin 1976a, b, Gauch, Whittaker & Wentworth 1977, Whittaker & Gauch 1978).

In the present paper we wish to comment on the investigations of the performances of the ordination methods based on the use of simulated coenoclines.

Coenocline simulation

A coenocline may be defined as a directed sequence of given species responses along a gradient (Fig. 1). The simulation of coenoclines, sometimes extended to more than one gradient, known as a coenoplane (Austin & Noy-Meir 1971, Gauch & Whittaker 1976), is often based on the assumption that the species responses (score) along a gradient may be accurately described as a bell-shaped curve, frequently of the Gaussian type. The simulation of a coenocline was recently used by Kessel & Whittaker (1976) to estimate the interaction of noise and beta diversity (cf. Whittaker 1972b) in ordination methods, including discriminant function analysis, Bray & Curtis ordination, and principal component analysis (PCA). Austin (1976b) has

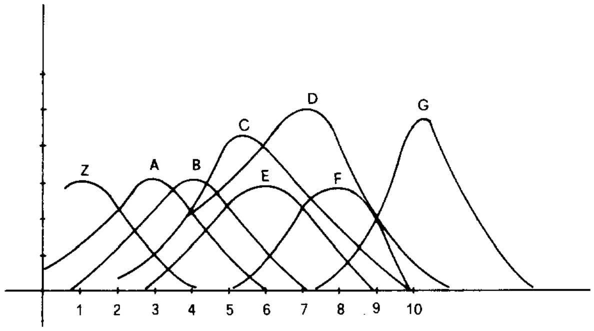

Fig. 1. A simulated coenocline with eight species (Z, A, B, C, D, E, F, G). The numbers on the gradient axis represent the points of ten hypothetical relevés. The scores of the species in these relevés are given in Table 1.

---

* The research was supported by a grant from the Italian CNR (Programma Finalizzato 'Promozione Qualità dell'Ambiente – sottoprogetto Metodologie Matematiche e Basi di Dati').

** We would like to thank dr. M. Dale, dr. E. van der Maarel and prof. L. Orlóci for useful comments on the manuscript and Mr. A. Zampar for technical assistance.

compared reciprocal averaging (Hill 1973, cf. also Hill & Smith 1976), Gaussian ordination (Gauch, Chase & Whittaker 1974, Ihm & van Groenewoud 1976, Van Groenewoud 1976), parametric mapping (cf. Noy-Meir 1974), and non-metric multidimensional scaling (Shepard & Carrol 1966, Kruskal 1964a,b, Anderson 1971), in terms of their sensitivity to different curvilinear responses of the species within simulated coenoclines.

The problem of linearity seems to play a fundamental role in these attempts for the evaluation of the ordination results. The basic objection against the application of linear methods (PCA and related methods) lies in the fact that the non-linear correlations produce distortion and loss of information in the ordinations (type A distortion, Orlóci 1974). However, before proceeding any further, it may be useful to clearly distinguish between linear and non-linear models. We prefer to define as linear all the models based on linear matrix algebra, or those approximating this (cf. van der Maarel 1969). It is noted that the data transformations (non-centering, centering, double centering, standardization, normalization and so on; Orlóci 1967, 1978a, Noy-Meir 1973 Noy-Meir & al. 1975, Feoli 1977) has been shown to lead to different performances of the linear models, and a very rich terminology, e.g. centered PCA, non-centered PCA, principal coordinate analysis (Gower 1966), reciprocal averaging (Hill 1973, Hill & Smith 1976).

With the term linear ordination we designate all methods involving the extraction of the eigenvalues (λ) and eigenvectors (α) of resemblance matrices. In this context it is preferable to define the ordination method on the basis of the resemblance function used to obtain the symmetric matrix. So, the ordination produced by the eigenvectors of a resemblance matrix of, for example, Sørensen's index will be referred to as a linear ordination based on Sørensen's index, and so on. PCA is a special case of linear ordination based on covariance (correlation coefficient, euclidean distance). It has the property of reproducing exactly the relative positions of the points in a metric space. The two examples of Fig. 2a,b and the examples given by Orlóci (1978) illustrate graphically such a property. Nichols (1977) presented an analytical discussion. When PCA is applied to the relevé of a coenocline, PCA re-establishes the relevés, as points, in a metric space, and if the relevé points in an ordination scatter diagram have a horseshoe type dispersion, this will mean that in the metric space the relevé points do in fact have such a dispersion. The requirement that a coenocline should be always represented by a straight line in a PCA ordination is unrealistic. There is of course a good reason to prefer another method: in PCA two or more axes may be necessary to reproduce an essentially unidirection trend, associated with a dispersion of relevés along a coenocline, instead of only one (Orlóci 1979). Feoli & Feoli Chiapella (1976), Feoli Chiapella & Feoli (1977) and Werger (1978) have identified gradients by interpolation based on species or relevé points in ordinations produced by PCA. Phillips (1978) suggested polynomial regression of principal components to reveal gradients (ordination axes). These are examples of the profitable use of PCA in indirect gradient analysis. The fact that the interpretability of the axes is difficult, and the statistical use of PCA in most cases is unrealistic is well known in phytosociology (Orlóci 1978a).

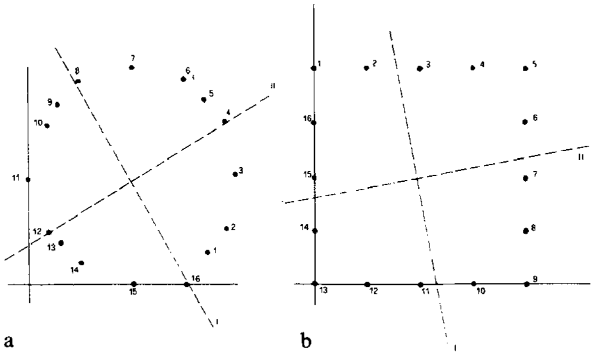

Fig. 2. Ordination by Orlóci's method (1966) of a set of points arranged on a circle (a) and on a perimeter of a square (b). The co-ordinates of the points are given in Table 2a, b. The dotted lines are the first two principal components.

Coenoclines and resemblance of relevés

A set of relevés may be related to one or several coenoclines. Each relevé may be described by a set of abiotic variables and by a set of species. In such a setup, the ecological problem amounts to find the correlation between individuals or sets of abiotic variables and the vegetation. This can be done by using different graphic methods, i.e., by superimposing on ordination diagrams, based on the species, the scores of abiotic variables in the corresponding relevés (Gittins 1965) or using numerical methods, both parametric and non-parametric ones (cf. Gounot 1969, Orlóci 1972a, Stanek, Jeglum & Orlóci 1977, Feoli 1976).

The resemblances of relevés may be calculated by different coefficients (Orlóci 1972b, 1978a, Goodall 1978) in two basically different ways: (a) considering the sets of abiotic and biotic variables jointly and (b) considering them separated. Recently, in his discussion on the covariation of plant species along ecological gradients, van Groenewoud (1976) stated that since the responses of species to a gradient are conceived as bell-shaped curves, the covariance, product moment correlation coefficient (*r*) and euclidean distance have little ecological value. This supposition could be considered true only if we consider an analysis of the relevés based on the joint use of biotic and abiotic variables.

However, if only the species are used as relevé descriptors, the argument of van Groenewoud becomes much weaker, and his interpretation of the correlation coefficient seems to be too restricted. It is true that the use of the product moment correlation coefficient in statistical tests of covariation is conditional on the assumption that the set of points:

$$G_R = \{(x, y) : x \in X \wedge y \in Y, xRy\} \qquad (1)$$

is such that the Cartesian product of two sets of measures on the variables *X* and *Y* will fit to a straight line, it is also true that the points may be so dispersed around this line that the correlation coefficient has no statistical meaning. Furthermore, the fact that the points (x, y) in G_R are lying on opposite sides of a straight line (see van Goenewoud 1976) may have a definite ecological meaning it is not a problem to be blamed on the mere application of the product moment correlation coefficient.

Since in an ecosystem different species do not tend to occupy the same niche (cf. Volterra 1931, Lotka 1932, Whittaker 1972b among others), it is obvious that for two species the same bell-shaped curve cannot be expected. The areas delimited by the curves should be more or less overlapping but not coincident. The problem to establish if *r* is a really inefficient measure of covariation, may be faced in the context of a simulated coenocline. The coenocline of Fig. 1 and Table 1 and the one presented by Austin (1976a, Fig. 4) are quite suitable for this purpose. The relationship of *r* and the relative overlapping areas of the bell-shaped curves of the species is important. The relative overlap (*oR*) of two species may be calculated by the following formula:

$$oR_{ij} = \frac{S_{ij}}{S_i + S_j - S_{ij}} \qquad (2)$$

where S_i is the integral of the bell-shaped curve of species *i*, S_j is the integral of the bell-shaped curve of species *j*, and S_{ij} is the overlap area. The integrals in this case have been calculated by a planimeter. Formula (2) is analogous to the similarity ratio coefficient (cf. Westhoff & van der Maarel 1978). Fig. 3 gives the graphics of regression for *r* on *oR* based on the data of coenoclines presented in Fig. 1 and by Austin (1976a, Fig. 4). The regression is clearly linear, meaning that the deterministic use of *r* should not be too problematic in gradient analysis.

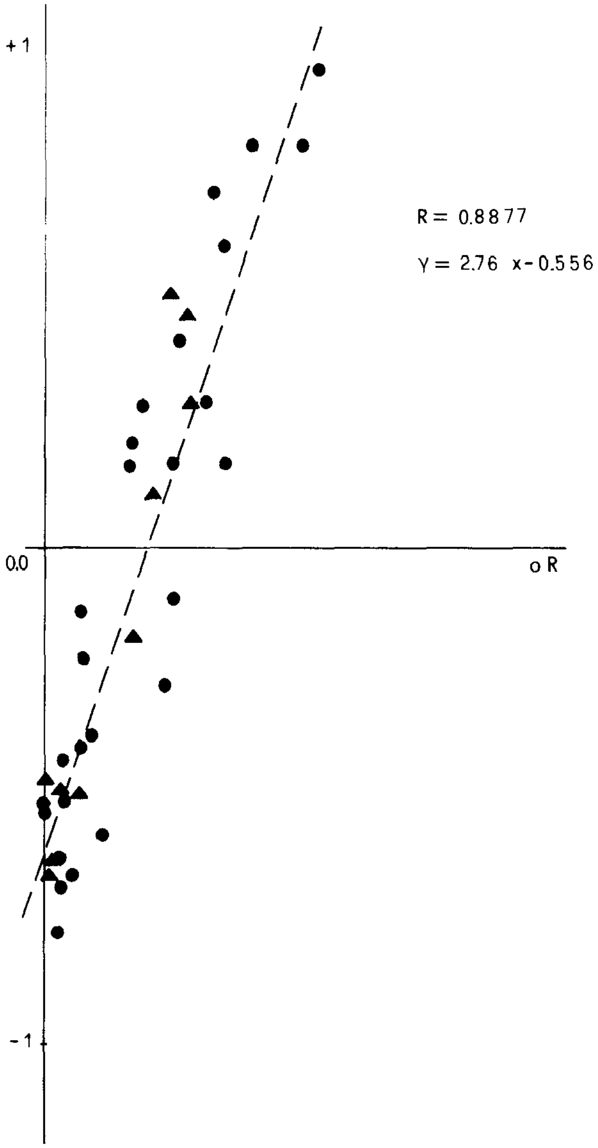

Fig. 3. Regression of the product moment correlation coefficient (*r*) of species on the relative overlapping area (*oR*) computed according to formula (1). Legend to symbols: ▲ data of coenocline presented by Austin (1976a, Fig. 4); ● data of coenocline of Fig. 1; *R* = correlation coefficient.

Table 1. Scores of the species in ten hypothetical relevés of simulated coenocline presented in Fig. 1. The numbers identify relevés, the letters identify species.

	1	2	3	4	5	6	7	8	9	10
Z	3	2	1	0	0	0	0	0	0	0
A	1	2	3	2	1	0	0	0	0	0
B	0	1	2	3	2	1	0	0	0	0
C	0	0	1	2	4	4	3	2	1	0
D	0	0	0	2	3	4	5	4	2	0
E	0	0	0	1	2	3	2	1	0	0
F	0	0	0	0	0	1	2	3	2	1
G	0	0	0	0	0	0	0	1	2	3

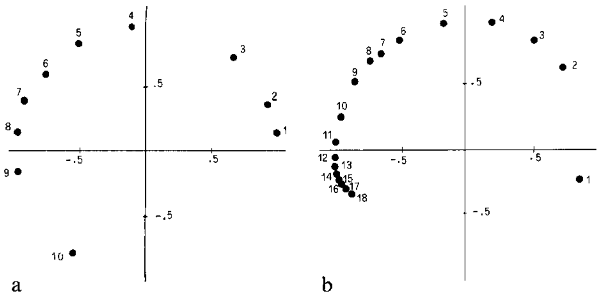

Fig. 4. a) Ordination of the coenocline of Fig. 1. b) ordination of the coenocline presented by Austin (1976a, Fig. 4). In both cases the axes are given by the method proposed by Feoli Chiapella & Feoli (1977).

Gradient analysis by a linear method

Feoli Chiapella & Feoli (1977) proposed a method of ordering relevés based on PCA applied to the log-similarity ratio matrix between them. The log-similarity between the relevés constitutes a special data transformation (see van der Maarel 1979 for other transformations of phytosociological data) which in PCA tends to give very regular disposition of relevés along arcs. As shown by van der Aart & Smeenk-Enserink (1975) the simple log-transformation of the normalized data has the same effect, however by using the log-similarity ratio each relevé is redescribed based on new data i.e. the similarities with other relevés. Although these data are not independent, they allow to apply more correctly the product moment correlation coefficient to measure the similarity between the relevés, since each relevé is now described by the same variable: similarity. The application of the product moment correlation coefficient produces a hyperspherical space where the relevés are settled on the surface of the hypersphere. If in a data set there is a dominant trend which can be explained by a gradient or by a compositional gradient the relevés will be arranged on an arc or a more or less twisted line. If no dominant trend is present, then the relevés will be arranged on the surface of the hypersphere in more or less dense clouds.

The application of Feoli Chiapella & Feoli's (1977) method to the simulated coenocline of Fig. 1 and of Fig. 4 of Austin (1976a) has led to ordination patterns in which

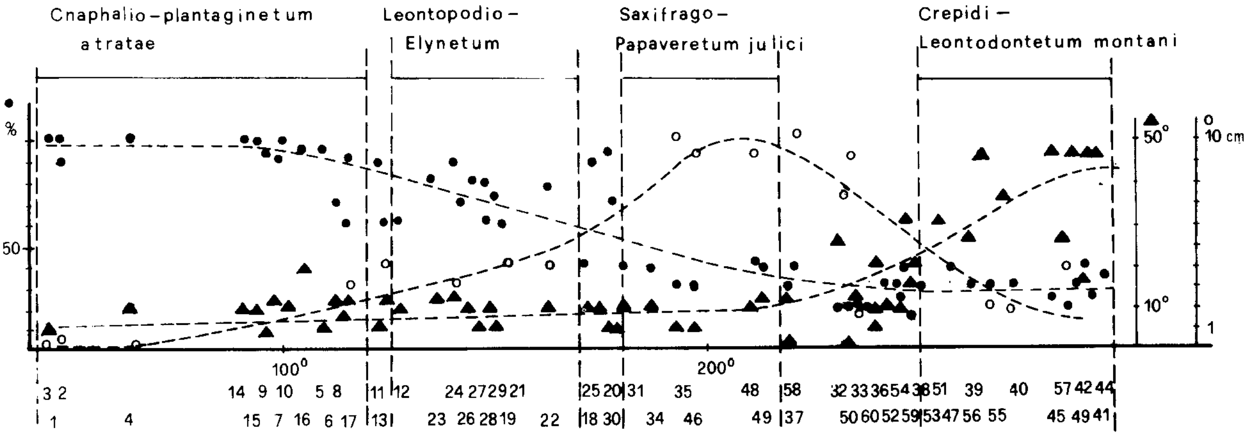

Fig. 5. Seriation of relevés from the summit of the Majella massive (Feoli Chiapella & Feoli 1977, Fig. 5) according to the angle of their position vectors given by the first two principal components (see the text for explanation). Legend to symbols: ○ = mean dimension of stones in each relevé; ▲ = slope angle; ● = percentage of ground surface covered by vegetation. The numbers along the abscissa identify the relevés as in Tables 1, 2, 3, 4 of Feoli Chiapella & Feoli (1977).

the fit of the first two axes to the arc is almost perfect (Fig. 4a, b). Along the arcs the rank order of the relevés is quite correct. In this case, the linear ordination has given the same results as a non-linear ordination (Austin 1976b, Fig. 4), and the preference for one or another is only a matter of taste. When the arrangement of the relevés fits to an arc, two axes can account for the trend of variation and the angles that the relevé vectors enclose can be used to arrange the relevés on a straight line. The angles have been already used to seriate sites by van der Aart & Smeenk-Enserink (1975).

The problem arises how to choose the starting point. This can be chosen indifferently among one of the two extreme points of the arc fitting the relevé points in the ordination scatter diagram. An example of such a seriation is shown in Fig. 5 for the relevés of the Majella's summit (Feoli Chiapella & Feoli 1977). In this way curved trends can be simply unfolded on straight lines. Such a seriation allows us to give a clear graphical description of the relationships between the four associations and the surface area of ground covered by the vegetation, the slope angle, and the mean dimension of stones. Furthermore, the seriation of the relevés can be useful to quantify the degree of overlap between the associations along the gradient. From Fig. 5, by considering the segments where the relevés belonging to two successive associations are mixed, it appears that the associations with the highest degree of overlap are *Saxifrago-Papaveretum julici* and *Crepidi-Leontodontetum montani*. This observation corresponds with the results of cluster analysis which show that the similarity between these two associations is higher than the similarity between *Gnaphalio-Plantaginetum atratae* and *Leontopodio-Elynetum* (cf. Feoli Chiapella & Feoli 1977, Fig. 4).

Table 2. Co-ordinates of ten points arranged on a circle (a) and on a perimeter of a square (b).

A

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
17	19	20	19	17	15	10	5	3	2	0	2	3	5	10	15
3	5	10	15	17	19	20	19	17	15	10	5	4	2	0	0

B

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	5	10	15	20	20	20	20	20	15	10	5	0	0	0	0
20	20	20	20	20	15	10	5	0	0	0	0	0	5	10	15

Conclusions

What Thurstone (1947) says in his introduction to Factor analysis should be considered by many data analysts:

'If the scientist takes his numerical coefficients very seriously at the exploratory stage he may be lacking in a desirable sense of humor about the crudeness of all his tools in spite of their polished appearance'.

The application of multidimensional scaling to plant ecology may be considered to be still at an exploratory stage since it is very difficult to postulate a general species response pattern to gradients (Austin 1972, 1976a). However the application of linear methods has always given positive results for those who have a proper understanding of them and consequently interpreted the results. The knowledge that linear methods arrange coenoclines in an involute horseshoe-shaped curves can indeed help the ecologist to use the results in a profitable way. The basic misunderstanding of those who on the basis of simulated coenoclines suggest the inadequacy of linear ordinations in gradient analysis relies on the fact that they have not considered some of the intrinsic aspects.

Summary

The evaluation of ordination methods through simulated coenoclines is discussed. The inconsistency of some arguments against the use of linear ordinations is stressed. The regression between the product moment correlation coefficient and the relative overlapping area of the bell-shaped curves of the species of coenoclines may be fitted to a linear function. An example of a seriation of relevés along a straight line based on a linear method is given using field data.

References

Aart, P.J.M. van der & N. Smeenk-Enserink. 1975. Correlation between distributions of hunting spiders (Lycosidae, Ctenidae) and environmental characteristics in a dune area. Neth. J. Zool. 25: 1–45.

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

Austin, M.P. 1972. Models and analysis of descriptive vegetation data. In: J.N.R. Jeffers (ed.): Mathematical models in ecology, p. 61–86. Blackwell, Oxford.

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. & I. Noy-Meir. 1971. The problem of non-linearity in ordination: experiments with two gradient models. J.Ecol. 59: 762–773.

Beals, E.W. 1973. Ordination: mathematical elegance and ecological naïveté. J.Ecol. 61: 23–35.

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

Feoli, E. 1976. Correlation between single ecological variables and vegetation by means of cluster analysis. Not. Fitosoc. 12: 77–82.

Feoli, E. 1977. On the resolving power of principal component analysis in plant community ordination. Vegetatio 33: 119–125.

Feoli, E. & L.Feoli Chiapella. 1976. Due associazioni rupicole della Majella. Not. Fitosoc. 12: 67–75.

Feoli Chiapella L. & E. Feoli. 1977. A numerical phytosociological study of the summits of the Majella massive (Italy). Vegetatio 34: 21–39.

Gauch, H.G.Jr. & R.H. Whittaker. 1972a. Coenocline simulation. Ecology 53: 446–451.

Gauch, H.G.Jr. & R.H. Whittaker. 1972b. Comparison of ordination techniques. Ecology 53: 868–875.

Gauch, H.G.Jr. & R.H. Whittaker. 1976. Simulation of community patterns. Vegetatio 33: 13–16.

Gauch, H.G.Jr., G.B. Chase & R.H. Whittaker. 1974. Ordination of vegetation samples by Gaussian species distribution. Ecology 55: 1382–1390.

Gauch, H.G.Jr., R.H. Whittaker & R.T. Wentworth. 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. I. A stand ordination. J.Ecol. 53: 385–401.

Goodall, D.W. 1954. Objective methods for the classification of vegetation. III. An essay on the use of factor analysis. Austral.J.Bot. 2: 304–324.

Goodall, D.W. 1978. Sample similarity and species correlation. 2nd ed. In: R.H. Whittaker (ed.): Ordination of plant communities, p. 99–149. Junk, The Hague.

Gounot, M. 1969. Méthodes d'étude quantitative de la végétation. Masson, Paris, 314 pp.

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

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

Hill, M.O. 1973. Reciprocal averaging: an eigenvector method of ordination. J.Ecol. 61: 237–249.

Hill, M.O. & A.J.E. Smith. 1976. Principal component analysis of taxonomic data with multi-state discrete characters. Taxon 25: 249–255.

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

Kessel, S.R. & R.H. Whittaker. 1976. Comparison of three ordination techniques. Vegetatio 32: 21–29.

Kruskal, J.B. 1964a. Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. Psychometrika 29: 1–27.

Kruskal, J.B. 1964b. Non-metric multidimensional scaling: a numerical method. Psychometrika 29: 115–129.

Lotka, A.J. 1932. The growth of mixed populations: two species competing for a common food supply. J. Washington Acad. Sci. 22: 461–469.

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

Maarel, E. van der. 1971. Basic problems and methods in phytosociology. Rep. 14th Int. Symp. Rinteln 1970. Vegetatio 22: 269–273.

Maarel, E. van der. 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity. Vegetatio 39: 97–114.

Nichols, S. 1977. On the interpretation of principal component analysis in ecological contexts. Vegetatio 34: 191–197.

Noy-Meir, I. 1973. Data transformation in ecological ordination. I. Some advantage 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. & M.P. Austin. 1970. Principal component ordination and simulated vegetation data. Ecology 51: 551–552.

Noy-Meir, I., D. Walker & W.T. Williams. 1975. Data transformation in ecological ordination. II. On the meaning of data standardization. J.Ecol. 63: 779–800.

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

Orlóci, L. 1966. Geometric models in ecology. I. The theory and application of some ordination methods. J.Ecol. 54: 193–215.

Orlóci, L. 1967. Data centering: a review and evaluation with reference to component analysis. Syst. Zool. 16: 208–212.

Orlóci, L. 1972a. On information analysis in Phytosociology. In: E. van der Maarel & R. Tüxen (eds.), Grundfragen und Methoden der Pflanzensoziologie, Ber. Int. Symp. Rinteln 1970, p. 75–88. Junk, Den Haag.

Orlóci, L. 1972b. On objective functions of phytosociological resemblance. The American Mid. Natur. 88: 28–55.

Orlóci, L. 1974. On information flow in ordination. Vegetatio 29: 11–16.

Orlóci, L. 1978a. Multivariate analysis in vegetation research. Junk, The Hague, 2nd ed. 451 pp.

Orlóci, L. 1978b. Ordination by resemblance matrices. 2nd ed. In: R.H. Whittaker (ed.) Ordination of plant communities, p. 239–275. Junk, The Hague.

Orlóci, L. 1979. Non-linear data structures and their description. International Statistical Ecology Program. Parma 1978.

Phillips, D.L. 1978. Polynomial ordination: field and computer simulation testing of a new method. Vegetatio 37: 129–140.

Shepard, R.N. & J.D. Carrol. 1966. Parametric representation of non-linear data structure. In: P.R. Krishnaiah (ed.), Multivariate analysis, p. 516–592. Academic Press.

Swan, J.M.A. 1970. An examination of some ordination problems by use of simulated vegetation data. Ecology 51: 89–102.

Stanek, W., J.K.Jeglum & L.Orlóci. 1977. Comparisons of peatland types using macronutrient contents of peat. Vegetatio 33: 163–173.

Volterra, V. 1931. Variations and fluctuations of the number of individuals in animal species living together. In: Chapman (ed.), Animal ecology with special reference to insects, pp. 409–448. Mc Grow-Hill, New York.

Werger, M.J.A. 1978. Vegetation structure in the southern Kalahari. J. Ecol. 66: 933–941.

Westhoff, V. & E. van der Maarel. 1978. The Braun-Blanquet approach. 2nd ed. In: R.H. Whittaker (ed.), Classification of plant communities, Junk, The Hague. p. 287–399.

Whittaker, R.H. 1967. Gradient analysis of vegetation. Biol. Rev. 49: 201–264.

Whittaker, R.H. 1972a. Convergences of ordination and classification. In: E. van der Maarel & R. Tüxen (eds.), Grundfragen und Methoden der Pflanzensoziologie, Ber. Int. Symp. Rinteln 1970. p. 39–57.

Whittaker, R.H. 1972b. Evolution and measurement of species diversity. Taxon 21: 213–251.

Whittaker, R.H. & H.G. Gauch. 1978. Evaluation of ordination techniques. 2nd ed. In: R.H. Whittaker (ed.), Ordination of plant communities, p. 277–336. Junk, The Hague.

Accepted 4 December 1979