Journal of Vegetation Science 3: 119-124, 1992

How reliable are our vegetation analyses?

Lepš, Jan[1] & Hadincová, Věra[2]

[1]*Biological Faculty, University of South Bohemia & Department of Biomathematics, Czechoslovak Academy of Sciences, Branišovská 31, CS-370 05 České Budějovice, Czechoslovakia; Tel. +42 38 817 ext. 214; E-mail KRIVAN@CSEARN;* [2]*Botanical Institute, Czechoslovak Academy of Sciences, CS-252 43 Průhonice u Prahy, Czechoslovakia; Tel. +42 2 750393*

Abstract. Two sets of 40 relevés, made independently by two observers on the same 5m × 5m sample plots, were compared to estimate the sampling error and to assess the effect of this sampling error on (1) estimates of species richness and diversity (2) results of multivariate analyses, and (3) estimation of species turnover in repeated sampling. The relevés were made according to the standard Braun-Blanquet method. The sampling error was estimated for (1) recording of species in sample plots and (2) visual estimation of the degree of cover (or of the general population size). Despite the fact that the sample plots were searched thoroughly for 30 - 40 min, the number of overlooked species was high with a discrepancy of 13% between corresponding relevés.

Regarding multivariate analysis, the error caused by missing species was at least as important as the error in visual estimation of species cover. The estimates of degree of cover using the Braun-Blanquet scale are sufficiently reliable for use in multivariate analysis when they are subjected to ordinal transformation. When average cover values are used, the patterns detected are based solely on dominants. Species richness and species diversity could be reliably estimated from the relevés, but the estimates of equitability are very unreliable. The classical relevé method remains one of the most efficient survey methods for recognition of vegetation types on the macro-community and landscape scales.

Keywords: Classification; Efficiency; Multivariate analysis; Ordination; Pseudo-turnover; Sampling error; Visual estimate.

Introduction

The most common description of a plant community in phytosociology is that based on the Braun-Blanquet relevé method (e.g. Mueller-Dombois & Ellenberg 1974; van der Maarel 1979). A relevé is a list of all species found in a sample plot with semi-quantitative visual estimates of abundance and cover, usually according to some scale (the Braun-Blanquet scale proper or a similar scale such as the Domin scale). Basically, the relevé should include all vascular plants, and also bryophytes and lichens; in practice reliable results are obtained with the vascular plants only (e.g. Krahulec, Rosén & van der Maarel 1986). The Braun-Blanquet approach is considered efficient for surveys, particularly because it is very fast. At present, there are probably over 100 000 relevés available, published or/and stored in data bases. No other sampling method has been (nor likely will be) able to provide such an extensive documentation of a wide range of vegetation types.

Nevertheless, there are two problems involved in the use of relevés in an ecological synthesis, particularly in numerical analyses. 1. Sample plots are usually selected intentionally, as 'typical' or 'representative' for some community type; hence they are not random samples. 2. The visual estimates of abundance and cover are subjective. This paper deals with the second problem.

The subjectivity in the estimation of plant species performance (cover, abundance) has been criticized from the early days of quantitative ecology (see Greig-Smith 1957; Kershaw 1964), and earlier authors (e.g. Hope-Simpson 1940) were also aware of errors in subjective estimates. This criticism includes, whether implicitly or explicitly, that presence-absence data are more reliable, as they do not contain the error of visual estimation. However, Tüxen (1972), who compared relevés of the same plot made simultaneously by different investigators, showed that the various species lists differed considerably: "It is meaningless to want to work quantitatively if one does not know for sure the degree of precision of the methods used and that of the yielded material." (translated from German). Dutch phytosociologists, comparing their cover % estimates, both in real vegetation and on models, found a large variation, especially for very low and very high cover values, but observed that most of the variation was eliminated after transformation to Braun-Blanquet scale values (van der Maarel, pers. comm.).

Visual estimates are not only used in classical phytosociology. Cover is often estimated directly for other purposes, usually as a percentage. A few comparative studies have been done for such direct estimates of cover (e.g. Sykes, Horrill & Mountford 1983). There are also some studies on errors in species inventories (e.g. Nilsson & Nilsson 1985; Kirby et al. 1986). Nilsson & Nilsson (1985), comparing the species composition of whole islands as recorded by two independent teams, calculated a 'pseudo-turnover' of species, to be taken into account when judging immigration and extinction rates. Rusch & van der Maarel (in press) used differences between repeated observations of small permanent plots (10 cm × 10 cm) for estimating 'spurious turnover' to which observed year-to-year turnover could be compared.

One source of error in vegetation analysis, which is probably underestimated, is the misidentification of species. In published phytosociological tables, only few species appear as unidentified or identified with uncertainty. Nevertheless, we often have problems, particularly with young individuals, even of species one is familiar with (cf. Tüxen 1972; Clymo 1980).

Observation errors will affect the results of multivariate analysis and calculations of community diversity. Gotfryd & Hansell (1985) have shown this for Principal Component Analysis (although in their case, a standardized sampling method was used).

In any type of vegetation research, one should be aware of possible errors that could be introduced in various ways. The aim of this paper is to estimate the size of the error that is introduced into the data by visual estimation of 'the degree of cover' and by the overlooking or misidentification of species, and to discuss its implications for further data treatment.

Material and Methods

Sampling

We analyzed 40 plots (most of them 5 m × 5 m) in the upper part of a watershed of the Rolava River in the Krusné hory Mts. in north-west Bohemia, Czechoslovakia (at 800 to 1000 m a.s.l.). The usual Braun-Blanquet scale was used: *r*: extremely rare, +: sparse, cover very small, 1: cover < 5%, 2: cover 5% - 25%, 3: 25% - 50%, 4: 50% - 75%, 5: 75% - 100%. 26 plots were abandoned oligotrophic meadows, ranging from dry to wet and not mown for 30 - 40 yr. We included wet meadows dominated by sedges (*Caricetalia fuscae* Koch 1926, *Scheuchzerietalia palustris* Nordhagen 1937), moist meadows with *Alopecurus pratensis* (*Arrhenatheretalia* Pawlowski 1925), and relatively dry stands with *Nardus stricta* (*Nardetalia* Preising 1949). The meadows were analyzed in July 1989. 11 plots represented communities on young clearcuts, and the remaining three were on disturbed peat-bog; they were all analyzed in August 1990. Each plot was analyzed by each author for max. 40 min. The authors worked independently, but simultaneously, while they could see each other. However, they did not consult each other about their identification problems, though they had the chance to consult one and the same specialist at the end of each field day. They did not adjust their analyses after consultation with each other. Only vascular plants were analyzed. 406 species are known from the area of 38 km^2 (Michálek, pers. comm.). About half of the species are restricted to ruderal places. In our plots we found 127 species. The authors were familiar with the flora of the area. Neither of them is a classical phytosociologist, but each uses relevés in research work and makes ca. 50 to 100 relevés per year. They do not belong to the same working group, usually do not make relevés together and there was no calibration or standardization of estimation before the sampling started. However, the authors' technique will have undergone a general 'calibration' during their training in the Braun-Blanquet approach, which is very common in Czechoslovakia.

Data analysis

First, the pairs of corresponding relevés were compared. Two types of inconsistencies were distinguished: differences in species composition, caused by overlooking of species (called discrepancy here) and differences in estimates of degree of cover. The discrepancy was calculated according to the formula of Nilsson & Nilsson (1985) for pseudo-turnover

$$\frac{A+B}{S_A+S_B} \times 100$$

where A and B are the numbers of exclusive species for each observer and S_A and S_B are numbers of species found by corresponding observers.

Species richness and species diversity

The values of species richness (i.e. number of species) and species diversity calculated according to the Shannon formula (i.e. $H' = \sum -p_i \log p_i$, where p_i is the ratio of importance of a species to the sum of importance of all the species) based on relevés of each author were compared. As importance value, the data transformed to cover values (midpoints of cover intervals for particular degree) were used, i.e. degrees of *r*, +, 1, 2, 3, 4, 5 were converted to values of 0.1, 0.5, 3, 15, 37.5, 62.5, 87.5, respectively. It should be noted that it is meaningless to calculate the diversity from ordinally transformed data (sensu van der Maarel 1979, i.e. into a 1- 9 scale). Since differences in importance values of particular species are very small, the resulting diversity is influenced nearly entirely by the number of species (in our case, the correlation between $e^{H'}$ calculated on the basis of data subject to ordinal transformation, and number of species was 0.998). Consequently, we used transformation to cover % values. The antilogarithm of H' was used as an alternative diversity measure and its ratio to the number of species as a measure of equitability (if all the species are equally represented in a relevé, the antilogarithm of H' equals the number of species).

Multivariate analysis

The entire data set of 80 relevés and separate data sets from each author were subject to Detrended Correspondence Analysis (Hill & Gauch 1980) using the CANOCO program (ter Braak 1987), with detrending by second order polynomials. The data were subject to three types of transformation (van der Maarel 1979): to presence - absence, ordinal transformation (i.e. conversion to values 1 to 7), and cover values. In this way we tried to compare the error introduced by visual estimation of cover and the error introduced by overlooking species. For comparison we performed a PCA (on covariance matrix, no standardization on samples). The similarity of ordination results was evaluated by correlation of relevé scores on the first ordination axis.

The occurrence of systematic bias was tested by Canonical Correspondence Analysis, with authorship of the relevé as an explaining variable. Since the possible bias was much smaller than the heterogeneity of the material, the number of relevés was used as a (categorial, i.e coded as 39 0/1 variables) covariable. In this way, the ordination reflected only the differences between observers. Accordingly, restricted permutations were used in the Monte Carlo permutation test.

Similarly, both sets were subjected to TWINSPAN (Hill, Bunce & Shaw 1975; Hill 1979) classification in two versions, one using the presence-absence data, and the other with quantitative data using each degree of the scale as a separate 'pseudo-species cut level'. The classifications were compared with the method of Goodman & Kruskal (1954).

Results

Species composition

The average discrepancy between relevés of the two sets was rather high - 13%. The average number of species per relevé found by both authors was 18.6; the average number of species found by a single author was only 16.4 (the same for both authors); it means that on average, approximately two species were overlooked by each author in each relevé. However, most of the overlooked species were rare - in 83 % of the cases they were assigned '*r*' or '+' (by the other author). As to the cover-abundance estimates: in 57.5% of all cases with species found by both authors, the estimate was the same, in 39.5% the difference was 1 degree on the Braun-Blanquet scale, and in only 3% the difference was greater than 1 degree. There was no systematic over- or underestimation by one author as compared with the other one.

Species richness and species diversity

There was no significant systematic observer bias, neither in species richness, nor in diversity, nor in equitability (paired t-tests). The correlation coefficient between numbers of species in corresponding relevés was 0.96. The correlation between diversity values was slightly less – $r = 0.83$ for corresponding H' values and $r = 0.80$ for $e^{H'}$. However, equitability values were correlated much less – $r = 0.48$. Apparently, the species richness component is satisfactorily estimated but the quantitative component co-determining equitability is far less reliably estimated.

Ordination

There are only small differences within pairs of relevés corresponding to the same plot in the DCA (ordinal data transformation) of all 80 relevés (Fig. 1). Separate ordinations of the data sets from each author were similar as well. There was little difference between the two ordinations based on presence - absence data and on (ordinally transformed) quantitative data respectively: correlations between first ordination axes were 0.991 and 0.988 respectively for DCA and 0.891 and 0.914 for PCA. Note, that for PCA the correlation is slightly higher when quantitative data are used. The concordance of ordinations based on data transformed to cover % values was much lower: $r = 0.623$ for DCA and $r = 0.563$ for PCA. Here, the ordination results depend mostly on the proportion of dominants. This was demonstrated by comparing ordinations of data sets without species with '+' or '*r*' in each relevé (more than half of all entries!). The results were almost identical ($r = 0.996$).

As usual, PCA and DCA gave very different results. An average linkage classification based on the matrix of correlations of first ordination axes (with the orientation of axes changed to produce only positive correlations) reveals that the influence of a sampling error is negligible compared to the influence of the choice between ordination methods (Fig. 2).

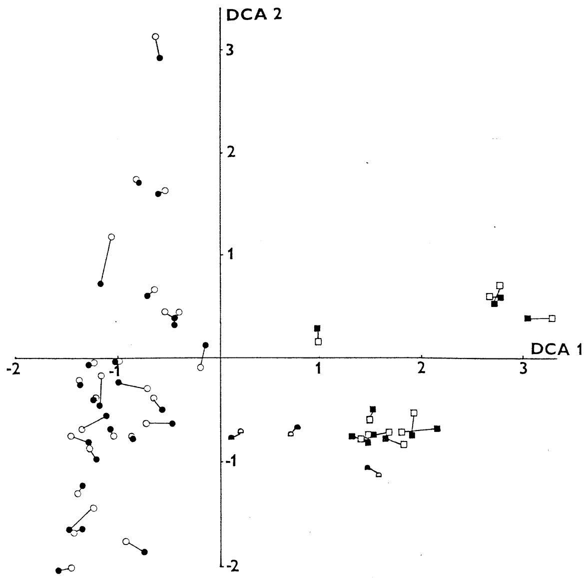

Fig. 1. Results of the detrended correspondence analysis of all 80 relevés. Relevés made by the first author are shown by closed symbols, relevés from the second author by open symbols. Circles, squares and semicircles correspond to meadows, clear cuts and regenerating peat bogs respectively. Relevés of the same plot are connected by a line.

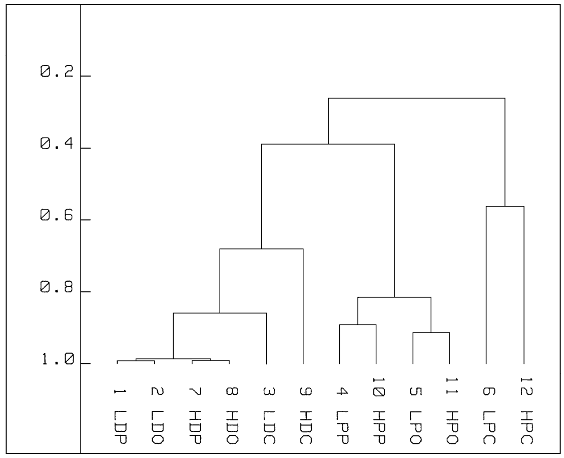

Fig. 2. Results of the classification of particular ordinations based on the correlation of the first axes. The first letter designates the observer, the second letter the ordination method (D stands for DCA, P for PCA), and the third letter data transformation (P - presence/absence, O - ordinal transformation, C - transformation to cover values).

With CCA no systematic bias could be detected. Relevé identity was used as a covariable and authorship as an explaining variable; consequently, the test is strong enough to detect systematic bias in any direction. The same analysis was carried out for data subjected to ordinal transformation and for presence-absence data. The eigenvalue corresponding to the first (constrained) axis was 0.017 (0.021 for presence - absence data), that for the second (unconstrained) one was 0.074 (0.076). The first one corresponds to the systematic bias, i.e. to systematic over- or underestimating of particular species by one of the authors, whereas the second one also reflects non-systematical differences between the two authors. Similarly, the Monte Carlo permutation test yielded low probability values: P = 0.55 - 0.45 for presence - absence. So, neither by direct comparison, nor with multivariate analysis could we detect systematic differences between relevés based on observer bias.

TWINSPAN classification

Four classifications were compared: for each author and for qualitative and quantitative data. Given the low number of relevés (40), only three division levels were compared. The first division was identical for all classifications; on the second level, three classifications were identical, the fourth differed only in one single relevé. The classification based on presence-absence data gave a different result. On the third hierarchical level, two relevés (out of 40) were placed into different groups in the corresponding classifications; one subgroup (out of four) was split in a completely different way (the same for quantitative and for presence-absence data). In both cases, the Goodman & Kruskal (1954) coefficient of classification similarity was 0.83. This means that the basic vegetation types were recognized quite similarly on the basis of the two data sets. It should be noted that even where the divisions were identical, the indicator species for particular divisions differed slightly.

Discussion

The discrepancy in species composition is high; it corresponds to the pseudo-turnover estimated by Nilsson & Nilsson (1985). It is lower than that in Tüxen's (1972) table, but the time we spent on the analysis was about twice as long. The discrepancy ('spurious turnover') for 10 cm × 10 cm plots in the data of Rusch & van der Maarel (1991) was lower, but still important. Clearly, all species censuses are affected by sampling error, regardless of the size of the plot searched. Discrepancy would probably have been lower if the moss layer had been analyzed, because while searching for mosses, one usually discovers some small specimens of vascular plants. Anyway, this value stresses the warning given by Nilsson & Nilsson (1985) for island-biogeographical studies (see also Raus 1988). In a repeated survey of permanent plots, where turnover is to be calculated, one usually has available results of the previous survey of the plot (island), and so the probability of overlooking the 'old' species is lower - however, thanks to this, the species number in some plots might increase simply due to the fact that 'old' species are usually remeasured and species overlooked in the previous survey are 'newly' found.

It is generally true that 'more time consuming methods are more accurate, but visual estimation is a common practice because it is rapid' (Sykes, Horrill & Mountford 1983), but in practice there are some constraints. Let us consider the point quadrat method as one more time consuming but possibly more accurate alternative. The greatest advantage of this method is not its precision but the fact that it is unbiased [though, see Goodall (1952) for the possible bias caused by the diameter of the pin]. If we consider pins randomly distributed in the area, the number of hits divided by the total number of pins is an estimate of the cover, and the number of hits is a binomial random variable with mean $= nc$ and variance $= nc(1-c)$, where n is total number of pins and c is the real (unknown) cover of the population, expressed as a decimal. Let us consider, as an example, a species with a cover of 3%. To get half of the estimates into the interval 1% - 5%, which roughly means that half of the estimates will be correct as to allocation to Braun-Blanquet degree 1, we need ca. 50 pins, with a 22 % chance that the species is missed by all the pins. In our experience, this would, in a meadow community, require about one hour's work by two people. The relative precision of the method increases with increasing cover of the species; to get a 50% chance of determining 'correctly' degree 2 for a species with 15% cover, we need only 10 pins, with a probability of 20 % that the species will be missed by all the pins.

Consequently, to get a complete species list, the area should be thoroughly searched again, assigning some arbitrary small value to the species not hit by a pin. Species with Braun-Blanquet degree 1, + or *r* make up 77% of all species in all our relevés. Unless one really has a great deal of time for analysis, the point quadrat method is not a realistic alternative to visual estimation for species with low cover. Similarly, Everson, Clarke & Everson (1990) concluded that only the most frequent species could be monitored with sufficient precision by point techniques.

The primary aim of making relevés is the recognition of basic vegetation patterns in a certain area. The results obtained with the numerical ordination and classification analysis appeared to be quite insensitive to sampling errors, unless transformation to cover values was applied. The patterns recognized on the basis of untransformed cover values usually correspond less well to the patterns recognized by classical phytosociology than patterns recognized on the basis of ordinal transformation (van der Maarel 1979; Kovář & Lepš 1986). When the ordinal transformation is used for PCA, the observer bias on the results of ordination is smaller than when the presence-absence data are used. This suggests that it is not necessarily correct to say that presence - absence data are less influenced by observer bias than visual cover estimation data. If one wishes to recover dominance patterns based on untransformed cover values, the species with low cover are unimportant and the sampling procedure should be refined as to the cover of the dominants. Here the point-quadrat method may be realistic.

It is generally supposed that observer bias will be influenced by various factors such as time of day, the expertise of the observer, etc. We do not have data to demonstrate this (and we have no experience in psychology to do this), but it seems that the number of overlooked species was higher when the concentration of the observer was disturbed. In some cases species were observed in the field but were not entered on the list. (Those errors are particularly troublesome when the overlooked species may have had a high cover.)

Differences in the ordinations caused by observer bias are generally smaller than those reported by Clymo (1980), who used a much more time consuming method of survey: sites were characterized by frequency in 12 continuous square quadrats, estimated time 5 h per site. However, direct comparison is difficult as he only gave one replicated site and the time reported by him includes all the fieldwork, not only the time used for direct species recording. Moreover, Clymo (1980) probably sampled a narrower vegetational range; this means that the same difference in a result of vegetation analysis would then cause greater discrepancy in the ordination diagram. Nevertheless, the comparison suggests the usefulness of the relevé as a rapid survey method. The thousands of available relevés contain a great deal of information. The present study suggests that the visual estimate of species cover should not be a major problem for the use of this information.

Our results are constrained by the character of the material analyzed. First, only two observers were compared - however, it is difficult to obtain experienced plant ecologists content to serve as 'Guinea-pigs' for a longer time (we tried to avoid the use of inexperienced students). Nevertheless, in studies where more observers were used one often found systematic bias (e.g. Smith 1944). Sykes, Horrill & Mountford (1983) even suggested that it would be possible to adjust the estimates for observer bias. However, those studies comprised smaller numbers of plots (some of them sampled repeatedly); this suggests that the stability of bias will probably decrease with the heterogeneity and/or size of the vegetation surveyed.

The second constraint is the heterogeneity of the vegetation types encountered in the study. Clearly, the more heterogeneous the material analyzed, the less important are the sampling errors. The heterogeneity of our material corresponds to that usually encountered in vegetation surveys.

Acknowledgements. This project forms part of a broader study of flora and vegetation of the upper part of the watershed of the Rolava River supported by the Regional Museum of Sokolov. We thank our colleagues, particularly Jiří Hadinec and Jaroslav Michálek for their cooperation and taxonomic advice during the field work, and Eddy van der Maarel and J. B. Wilson for constructive comments on the manuscript. We thank Carla Bossard for correcting our English.

References

Clymo, R. S. 1980. Preliminary survey of the peat-bog Hummel Knowe Moss using various numerical methods. *Vegetatio* 42: 129-148.

Everson, T.M., Clarke, G. P.Y. & Everson C.S. 1990. Precision in monitoring plant species composition in montane grasslands. *Vegetatio* 88: 135-141.

Goodall, D. W. 1952. Some considerations in the use of point quadrats for the analysis of vegetation. *Aust. J. Sci. Res.* 5: 1-41.

Goodman, L. A. & Kruskal, W. H. 1954. Measures of association for cross classifications. *J. Am. Stat. Ass.* 49: 732-764.

Gotfryd, A. & Hansell, I. C. 1985. The impact of observer bias on multivariate analyses of vegetation structure. *Oikos* 45: 223-234.

Greig-Smith, P. 1957. *Quantitative Plant Ecology.* Butterworths, London.

Hill, M. O. 1979. *TWINSPAN. A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes.* Ecology and Systematics, Cornell University, Ithaca.

Hill, M. O. & Gauch Jr., H. G. 1980. Detrended correspondence analysis: an improved ordination technique. *Vegetatio* 42: 47-58.

Hill, M. O., Bunce, R. C. H. & Shaw, M. W. 1975. Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland. *J. Ecol.* 63: 597-613.

Hope-Simpson, J. F. 1940. On the errors in the ordinary use of subjective frequency estimations in grassland. *J. Ecol.* 28: 193-209.

Kershaw, K. A. 1964. *Quantitative and Dynamic Ecology.* Arnold, London.

Kirby, K. J., Bines, T., Burn, A., MacIntosh, J., Pitkin, P. & Smith, I. 1986. Seasonal and observer differences in vascular plant records from British woodlands. *J. Ecol.*74: 123-131.

Kovář, P. & Lepš, J. 1986. Ruderal communities of the railway station Česká Třebová (Eastern Bohemia, Czechoslovakia) - remarks on the application of classical and numerical methods of classification. *Preslia* 58: 141-163.

Krahulec, F., Rosén, E. & van der Maarel, E. 1986. Preliminary classification and ecology of dry grassland communities on Ölands Stora Alvar (Sweden). *Nord. J. Bot.* 6: 797-809.

Mueller-Dombois, D. & Ellenberg, H. 1974. *Aims and Methods of Vegetation Ecology.* Wiley, New York.

Nilsson, I. N. & Nilsson, S. G. 1985. Experimental estimates of census efficiency and pseudo-turnover on islands: error trend and between-observer variation when recording vascular plants. *J. Ecol.* 73: 65-70.

Raus, T. 1988. Vascular plant colonization and vegetation development on sea-born volcanic islands in the Aegean (Greece). *Vegetatio* 77: 139-147.

Rusch, G. & van der Maarel, E. in press. Species turnover and seedling recruitment in limestone grasslands. *Oikos* 62.

Smith, A. D. 1944. A study of the reliability of range vegetation estimates. *Ecology* 25: 441-443.

Sykes, J. M., Horrill, A. D. & Mountford, M. D. 1983. Use of visual cover assessments as quantitative estimators of some British woodland taxa. *J. Ecol.* 71: 437-450.

ter Braak, C. J. F. 1987. *CANOCO - a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal component analysis and redundancy analysis.* TNO, Wageningen.

Tüxen, R. 1972. Kritische Bemerkungen zur Interpretation pflanzensoziologischer Tabellen. In: van der Maarel, E. & Tüxen, R. (eds.), *Grundfragen und Methoden in der Pflanzensoziologie*, pp. 168 - 182. Junk, Den Haag.

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

Received 7 August 1991;
Revision received 20 December 1991;
Accepted 3 January 1992.