

A new multivariate approach to studying temporal changes of vegetation

J. Podani^{1,*}, P. Csontos², J. Tamás³ and I. Miklós¹

¹Department of Plant Taxonomy and Ecology, L. Eötvös University, Pázmány P. s. 1/c, H-1117 Budapest, Hungary; ²MTA-ELTE Research Group in Theoretical Biology and Ecology, Pázmány P. s. 1/c, H-1117 Budapest, Hungary; ³Department of Botany, Hungarian Natural History Museum, P. O. Box 222, H-1476 Budapest, Hungary; *Author for correspondence (e-mail: podani@ludens.elte.hu; fax: +36 1 3812 188)

Received 27 April 2004; accepted in revised form 11 March 2005

Key words: Classification, Life forms, Manhattan metric, Mantel test, Matrix rearrangements, Nearest neighbors, Ordination, Procrustes method, Seed mass spectra

Abstract

We emphasize the necessity of a complex approach to evaluating vegetation change at various levels of abstraction. The analytical steps include comparisons at the data, derived variable, distance, ordination and classification levels. A variety of data randomization methods incorporated in testing the significance of changes in raw data are introduced and compared. It is shown that these are true alternatives to Procrustean comparisons, which offer an apparently unfortunate choice in the presence/absence case. We propose to evaluate nearest neighbor relationships among quadrats in a new method, called adjacency analysis, to detect temporal trends that may remain unrevealed, should our attention be paid to full distance structures only. As an illustration, compositional and structural changes in the rock grassland vegetation of the Sas-hegy Nature Reserve (Budapest, Hungary), intensively sampled by quadrats in 1977 and 2000, are evaluated. Permutation tests show that differences between the 2 years are much smaller than expected by chance alone. Such an overall stability in community structure, however, does not mean that minor aspects of vegetation pattern are invariant over the years. Changes in life form and seed mass spectra are explained by the fluctuation of hemicryptophytes and the slight but detectable expansion of annuals and woody species. Classification is slightly rearranged in time, with clearly detectable within-cluster changes, also depicted in ordination scattergrams.

Nomenclature follows Simon (2000).

Introduction

Temporal change of community composition and structure has long been a central issue in vegetation ecology. Information on successional processes may be obtained from various sources, such as aerial photographs, palynological analyses of soil cores or fine scale studies recording exact positional changes of individual plants. However, conventional quadrats or sample plots represent the most appropriate means to collect data for evaluating

compositional change. There are two different approaches to quadrat sampling. In the first, a given study area is selected and then sampled repeatedly over time. The second approach involves placing the quadrats in sites of different, preferably known age, at a given point of time, assuming that the sample thus obtained represents various stages of a temporal series (chronosequence or space-for-time substitution, Pickett 1989; Foster and Tilman 2000).

Surveys utilizing the resampling strategy may proceed in two ways. If the objective is to evaluate

the relative importance of spatial and temporal variation, several *replicate quadrats* are placed closely and permanently at the same site, and are monitored over time. This is the admissible strategy if the effect of some environmental impact is to be detected by community changes, because spatial effects can be factored out from the comparisons. Clarke (1993) and Philippi et al. (1998) list a variety of hypothesis testing procedures applicable to this situation. However, the fact that replicates are used implies that each site is considered homogeneous in space or that there are well-circumscribed, externally defined vegetation units whose interrelationships are assumed to be invariant in time. This is rarely the case in community ecology, however. When no *a priori* assumptions can be made on homogeneity and on the existence of vegetation types or other groups, replicate samples do not apply (or all quadrats are replicates in some sense) and the study objectives are usually other than impact assessment. Such studies intend to analyze natural vegetation changes such as progressive succession or temporal fluctuations, which allow rearrangements in the classification structure. Clearly, these problems require completely different analytical treatments.

In situations with no *a priori* assumptions on stand homogeneity, a fundamental requirement is to ensure correspondence between sampling units representing temporal replicates from the same site (via permanent quadrats, for example). We emphasize that comparisons between the corresponding quadrats must be done simultaneously at several levels of abstraction, each representing a conceptual space associated with a given step of vegetation analysis (cf. Podani 1992). These spatial representations include the raw data, variables derived from the data (e.g., diversity, life form or other frequency distributions), resemblance, classification and ordination. The possibilities for studying temporal changes in these spaces, we believe, are not yet exhausted completely, and most published results are confined to one or a few spatial aspects mentioned above. We suggest a complex strategy to evaluate vegetation changes for the simplest situation, when two sets of data, reflecting two points in time, are available. A crucial part of the methodology is hypothesis testing based on a wide variety of randomization techniques. Special attention is paid to the comparison of data randomization alternatives by showing that Procrustean approaches to

direct comparison of presence/absence data are inferior to randomization designs. A new method, *adjacency analysis* is introduced to evaluate temporal changes based on distance matrices coming from the different years. Comparisons of ordinations as well as classifications are complemented with randomization-based significance tests. We demonstrate that taking either part of the methodology alone provides only an incomplete picture on temporal processes.

The use of the strategy outlined above is illustrated by examining changes of rock grassland vegetation. The study object is a plant assemblage characteristic of dolomite substrates in the Carpathian Basin. Whereas quite a few ecological investigations have been done in these rock grasslands (e.g., Zólyomi 1958; Debreczy 1987; Csontos and Lőkös 1992; Bartha et al. 1998; Podani 1985; Podani et al. 2000), all are confined to a given year, and no objective analyses of temporal changes have been made. In this study, the temporal extent is expanded to two points of time, 23 years apart, to satisfy the minimum requirement for evaluating natural compositional changes. We use a relatively large number of plots allowing a spatial extension that is less common in succession studies.

Material

The study area is the Sas-hegy (Sas Hill) Nature Reserve situated within the city limits of Budapest, Hungary. The reserve is of 30 ha area separated from the surrounding built-in urban land by fence. It has been under protection and all management practices have been suspended since 1958. Sas-hegy is composed of three main outcrops, the highest reaching an elevation of 266 m a.s.l. The slopes, especially the south facing ones, are steep, with an inclination of 45° or more. The bedrock is Triassic dolomite, subject to intensive superficial erosion of mostly physical nature. Extensive bare rock surfaces appear on very steep slopes, whereas gentle slopes are covered by thin layers of rendzina. This is a soil type with brown or black surface horizons underlain by light gray to pale yellow material developed from soft, highly calcareous parent rocks. Annual precipitation is around 650 mm, with maxima in May and September. The minimum air temperature may be as low as -15 °C in the winter, and the daily maximum may reach 35 °C in the

summer. The slopes exposed to the south considerably warm up, especially in the summer, whereas northern slopes have a much cooler climate throughout the year (Draskovits and Kovács-Láng 1968). The physical and chemical characteristics of the rock surface and the weather conditions are combined into a peculiar microclimate of each habitat, thus providing shelter for several plants of southern Europe on south-facing slopes and for glacial relicts on northern slopes (Zólyomi 1942).

Sas-hegy is covered by a mosaic of natural, semi-natural and secondary plant assemblages. The present study focuses on the less disturbed natural stands, which cover about a third of the total area of the entire reserve. The natural vegetation is confined to the hilltops and the surrounding steep slopes. Many species that are mainly responsible for the physiognomic structure appear throughout the study area, including *Festuca pallens*, *Carex humilis*, *Bromus pannonicus*, *Seseli leucospermum*, *Stipa eriocalis* and *Asperula glauca*. The vegetation forms a continuum in space, the microclimatic and aspect differences are not manifested in sharp community boundaries. Apparent imbalance in spatial distribution due to clear habitat preferences is exhibited only by some less widespread species, such as *Sesleria sadleriana* (an endemic species to the Carpathian basin) or *Chrysopogon gryllus*. Detailed multivariate analyses based on data from 1977 (Podani 1998; Podani et al. 2000 and references therein) confirm these observations and nevertheless show that based on presence/absence data sample plots are arranged around three vegetational noda in the ordinations and classifications. Since there is no one-to-one correspondence between these noda and the four plant community types ("associations") described from the same area according to the principles of the Zürich-Montpellier phytosociological school (Zólyomi 1958; Debreczy 1966), we shall focus on global analyses of the entire vegetation and omit details related to the distinguishability or recognizability of these associations.

Methods

Sampling

The first sampling survey was made by the senior author from March to September, 1977. A total of

80 quadrats were placed in the field in a manner which may be considered as a mixture of preferential and regular designs. The strategy was preferential in a sense that areas obviously degraded or overgrown by naturalized ornamental shrubs (mostly *Syringa vulgaris*) were discarded. The author attempted to ensure equal representation of the less disturbed, i.e., natural and semi-natural areas by placing the plots evenly in the field. Each square plot was a nested system of 8 small sub-quadrats increasing in size, from 0.5×0.5 to 4×4 m². Percentage cover of vascular plant species was recorded. This sampling design provided data used in a wide variety of studies, including a complex multivariate experiment analyzing the relative effect of plot size and data type on classifications and ordinations (Podani 1989). Its results were considered in selecting the plot size and the data type to be used in the repeated sampling survey in 2000. We selected the largest quadrat size because the analyses indicated a trend towards stabilization of results along with increasing quadrat size. Of the four data types, three of them obtained by simplifying the cover values along a data transformation series, we chose the presence/absence form. This choice is supported by lower sampling effort, decreased sampling error and the observation that plot size effects are considerably smaller in case of presence/absence data than when cover scores are used. The quadrat positions were identified in the field using notes and maps made in 1977, with an accuracy estimated at about a few metres. The reestablished quadrats (semipermanent plots, *sensu* Persson 1980) were visited repeatedly by the first three authors from April to September, 2000, to record presence/absence scores.

Data analysis

The data pertaining to the two study years are first contrasted using simple statistics concerning the number of species, and life form and seed size frequency histograms.

Life form changes

The Raunkiaer's life form categorization of species follows Simon (2000). The categories detected include the annuals (Th), biennials (TH), hemi-cryptophytes (H), geophytes (G), chamaephytes

(Ch), nanophanerophytes (N), microphanerophytes (M) and meso-megaphanerophytes (MM). The total number of occurrences of species was determined for each category and a frequency histogram was compiled for both years. The two frequency distributions were compared by the χ^2 homogeneity test.

Seed mass spectra

Thousand seed mass (TSM) data of species were taken from the SEED database (Csontos 1998) and then converted into eight seed mass classes as follows: 1 = less than or equal to 0.20 g, 2 = 0.21–0.50 g, 3 = 0.51–1.00 g, 4 = 1.01–2.00 g, 5 = 2.01–4.00 g, 6 = 4.01–10.00 g, 7 = 10.01–50.00 g, 8 = above 50.00 g. There was no information for about 7% of the species in the SEED database. For these species, scores were estimated based on known data of congeners (see Csontos et al. 2004, for details) As a result, only a negligible total of 29 occurrences, pertaining to 3 species, had to be discarded from the data. For the seed mass spectrum calculations, species were weighted with their number of occurrences in the corresponding set of 80 quadrats. Seed mass spectra for the 2 years were statistically evaluated by χ^2 tests for homogeneity and for trend (Altman 1991).

The vegetation of the two study years is compared at various levels of abstraction representing different stages of a conventional multivariate community analysis: data matrix, resemblance matrix, classifications and ordinations. To ensure consistency of tests, and comparability of results, all presence/absence-based statistical analyses rely directly or indirectly on Manhattan (city block) distance (Podani 2000b), which is the absolute number of differences between the two objects (sample plots, matrices, etc.) compared. For data matrices \mathbf{X}_1 and \mathbf{X}_2 , it is given by $d_{12} = \sum_i \sum_j |x_{1ij} - x_{2ij}|$.

Data matrices

Statistical consistency is not the only argument in favour of using Manhattan distance at the data level. Procrustean superpositions were also advocated by some authors (Jackson 1995; Peres-Neto and Jackson 2001) as being potentially applicable to the same purpose. Considering the current state-of-the-art of computer technology, superpositioning two presence/absence data matrices, even if the number of axes is so many as 141 as in our case, would not pose computational difficulties. The real problem is that since the number of possible values, i.e., coordinates in the multidimensional space, is only two in the binary case for each axis, the operation of dilation involved in Procrustes analysis appears less logical than for continuous variables. An even more serious difficulty is that two presence/absence matrices can have zero Procrustean dissimilarity (i.e., zero sum of squares) even if they do not have a single species in common. This is illustrated with the two artificial data matrices of Table 1. The rotation operation involved in Procrustean comparisons can perfectly fit matrix \mathbf{A} onto \mathbf{B} , and vice versa, because every species in \mathbf{B} has a similarly distributed counterpart in matrix \mathbf{A} . It is obvious therefore that direct Procrustes analysis of data matrices gives no information whatsoever on compositional agreement. This is why we feel that Manhattan distance is a more logical choice when comparing two alternative data matrices as a statistic to be tested for significance.

These tests are performed via *matrix randomization*, commonly used in ecology for evaluating species co-occurrence patterns (Connor and Simberloff 1979; Gotelli and Entsminger 2001). With the first matrix remaining intact, the second matrix is randomized and their distance is calculated many number of times (10,000 in this study). Then, the probability that the actual distance comes from

Table 1. Artificial presence/absence data matrices illustrating that Procrustean comparisons produce zero sum of squares even if the two sets have no species in common.

Species	Matrix A								Matrix B								
1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
3	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
5	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0

a comparison involving a random matrix is estimated by counting the number of distances larger than the actual one. This facilitates testing the general null hypothesis that vegetation composition has not changed at the data level. The crucial part of this procedure is the randomization procedure itself, which may imply radically different null hypotheses that were not yet compared in the literature of numerical ecology. We used five variants, all of them considering the grand total (g , the number of occurrences in the original matrix) as a fixed value:

1. *Unconstrained randomness.* A null matrix of the same size as the original is modified by entering g presence scores (1s) completely randomly. This means that all 0 entries of the starting matrix have the same chance of being transformed into 1. No restrictions are imposed on row and column sums, so that both sets of marginal values approximate the uniform distribution. This implies a null situation and thus serves as a reference basis for comparison with other models.

2. *Constant richness.* Into the null matrix, g presence values are entered such that the column (site) totals of the data matrix are kept constant. This strategy implies that each site can maintain only a fixed number of species, whereas there is no restriction regarding the frequency of occurrences for any species. The consequence of this strategy is that the resulting species totals approximate the uniform distribution.

3. *Constant frequency.* The random matrix is generated as above, but with the row (species) totals fixed. Thus, each species occurs as frequently as in the original data, whereas original site totals are disregarded. As a result of randomization, the site totals approximate the uniform distribution.

4. *Full constraints.* The most constrained approach is to keep both site and species totals fixed. It means that species frequencies and site richness values remain the same as in the original data. Complete randomization with these serious constraints presents a real challenge in numerical ecology and has been unsolved until recently (cf. Gotelli and Graves 1996; Gotelli and Entsminger 2001). We used our sum of squares reduction algorithm (Miklós and Podani 2004), which starts from a random matrix containing arbitrary non-negative integers with the condition that the row and column totals are the same as the respective

totals in the original data matrix. An obvious advantage of this strategy against other randomization methods (Gotelli and Entsminger 2001) is the independence of final results from the starting matrix. This is not associated with equiprobability of resulting matrices, however. Therefore, the simulated matrices are further perturbed by a large number of swaps. We used a variant of the Metropolis – Hastings algorithm (Metropolis et al. 1953) suggested by Miklós and Podani (2004) which guarantees equidistribution of resulting matrices. Each simulated matrix was perturbed by 100,000 trial swaps.

5. *Probabilistic marginals with adjustment to fixed grand total.* Since the above strategy does not allow the slightest fluctuation in marginal totals, we propose a modification of the probabilistic simulation model described in Gilpin and Diamond (1982). Following the contingency table approach, the probability that $x_{ij} = 1$ in the random matrix is calculated from the respective marginal totals (column total c_j , and row total r_i) and the grand total, g , of the original matrix according to

$$p\{x_{ij} = 1\} = c_j r_i / g.$$

The random matrix is filled up value by value according to the above probabilities. On the average, the marginal totals as well as the grand total (g') of the random matrix approximate the actual totals but complete agreement is exceptional. In order to keep at least the grand total fixed, $g' - g$ 1s are randomly replaced by 0 if $g' > g$, whereas $g - g'$ 0s are replaced randomly by 1 if $g > g'$ for each simulated matrix.

Distance matrices

For each year, Manhattan distance matrices among quadrats and among species were computed. The hypothesis that the distance structure has not changed significantly after 23 years was examined by the *Mantel test* (Manly 1991) for the sites, as well as for the species. The method involves random permutations of the rows (and therefore the columns) of the second matrix, whereas the first remains intact. Then, the product moment correlation coefficient is computed formally based on the respective non-diagonal elements of the matrices (*matrix correlation*). This procedure was repeated 10,000 times to yield a sample distribution of correlation values which can

then be used to estimate the probability that the correlation for the original matrices arises also from the randomized design.

Temporal and spatial variation in the study area was also analyzed by a new procedure, *adjacency analysis*, a technique not to be confused with nearest neighbor clustering to which it is logically related. A 160×160 distance matrix for all the sample plots was computed using the Manhattan measure (plots from 1977 numbered 1–80, those from 2000 numbered 81–160). Then, for each plot we found its nearest neighbor by identifying the smallest value in the corresponding row of the distance matrix. Note that the nearest neighbor (NN) relationships are not necessarily symmetric and can be represented by a directed graph in which vertex j is connected to vertex k if quadrat k is the nearest neighbor of quadrat j . We distinguish among three possible outcomes of this search:

1. “*Spatiotemporal permanence*”: Quadrat j is most similar to the quadrat made in the same site in the other study year (i.e., to quadrat $j+80$ or $j-80$, which is a “temporal pair” in the terminology of Philippi et al. 1998). This is the situation we expect if vegetation change in the site is less substantial than the difference from any other sites, an observation interpreted as an indication of temporal stability and represented in the NN graph by a directed edge from j to k . This differs substantially from the next two situations in which temporal pairs are broken.
2. “*Spatial permanence*”: Quadrat j is most similar to quadrat k sampled elsewhere in the same year (j and $k < 81$ or j and $k > 80$, a “spatial pair” of Philippi et al. 1998). This situation arises when temporal floristic changes of a given quadrat supersede the minimum difference of that quadrat from another quadrat taken in the spatial continuum. Quadrat j therefore represents a case of spatial stability or permanence.
3. “*Spatiotemporal mutability*”: Quadrat j is most similar to some other quadrat k sampled in a different year ($j < 81$ and $k > 80$ or $j > 80$ and $k < 81$, “mixed or unrelated pair” of Philippi et al. 1998). This implies situations in which a quadrat j has lost its close relationships to its temporal pair as well as with all of its spatial pairs, implying mutability in both time and space.

After detecting the number of sites representing the three cases discussed above, the distance

matrix was randomly permuted 10,000 times to yield corresponding random figures that can be used for significance testing. The appearance of ties does not influence the test, because the number of ties is invariant for permutations. The only consequence is that tied sites have to be ignored when the outcomes are counted.

Average Manhattan distances for the same year and between quadrats pertaining to different years are calculated from the pooled distance matrix. The average of between year distances is then tested for significance based on 10,000 permutations of the original matrix. Note that within year averages are based on $80 \times 79/2$ distance values, whereas the average between the two years is derived from more than twice as many quadrat pairs, i.e., 80×80 comparisons.

Ordinations and classifications

The Manhattan distance matrices between quadrats served as input to ordination (principal coordinates analysis, PCoA) and clustering (complete linkage strategy). The use of PCoA with Manhattan metric has been shown to be the best combination to reduce potential horseshoe effects in the ordination (Podani and Miklós 2002). The efficiency of distance recovery in the dendrograms was measured by the cophenetic correlation coefficient (Sneath and Sokal 1973). The similarity of dendrograms for the 1977 and 2000 data was expressed as the matrix correlation of their ultrametric matrices expressing cluster membership divergence, a topological dendrogram descriptor (Podani and Dickinson 1984; Podani 2000a). The significance of this result was tested using a permutation technique: the second dendrogram was randomly relabeled and its correlation with the other calculated 5000 times. The two ordinations were compared by Procrustes analysis (see e.g., Schönemann and Carroll 1970; Gower 1975) based on the first two dimensions. Our *a priori* argument supporting this choice is that only the first two axes are depicted and interpreted in most community studies. In any case, involving more and more dimensions would gradually reduce the problem to the comparison of starting distance matrices. The resulting sum of squares statistic was tested for significance by a permutation technique. In order to be consistent with the dendrogram comparison, the points in the ordination of second year quadrats were randomly relabeled and Procrustes

analysis was performed 5000 times to yield a reference distribution of the statistic.

The pooled, 160×160 distance matrix of quadrats was analyzed by PCoA to identify temporal trajectories in the ordination plane, following Austin's (1977) pioneering work. The same matrix was also subjected to complete linkage clustering to see whether temporal pairs are classified together at a given number of clusters. A similar, joint clustering approach was used, for example, by Sanders et al. (2004) to measure the stability of a classification over time. The number of temporal pairs that appear in the same cluster is tested for significance using random permutations of the partition. In each permutation, the objects are relocated to another cluster randomly by keeping the original cluster sizes unchanged.

As a supplementary tool to adjacency analysis, cases of spatiotemporal and spatial permanence and spatiotemporal mutability are identified in the classification. This is important to evaluate whether temporal change of vegetation is manifested in the classification.

Computer programs

The classifications, ordinations, the Mantel tests, Procrustes analyses and dendrogram comparisons and associated tests were performed using the SYN-TAX program package (Podani 1993, 2001). The sum of squares reduction algorithm for matrix randomization is implemented in a program written in C (see Supplement to Miklós and Podani 2004). The chi-square statistics were calculated using InStat (1997). All other analyses (adjacency analysis, partition permutations, etc.) were computed using small FORTRAN routines written exclusively for the purpose.

Results

Basic statistics

The number of species detected in the entire study area in 1977 was 123, which increased to 139 by 2000. Most of the species (115) occurred in both study years, 8 appeared only in 1977 while 24 were found only in 2000. Whereas this change in species numbers cannot escape the eye of the ecologist, there is a remarkable similarity between the numbers of occurrences in the two study years. In 1977, we recorded $g_1 = 3185$ species presences in the 80

quadrats, whereas 23 years later the number of occurrences was $g_2 = 3209$. The mean species richness in a $4 \times 4 \text{ m}^2$ sample plot was thus around 40 in both years. (The difference of 24 between the two totals is so small that no formal statistical test is required to conclude: the species richness per unit area was fairly constant over the study years.) Most of the 32 species that appeared only in one year are usually accidental elements in the communities: their total number of occurrences was 111, representing a negligible proportion (1.7%) of the pooled grand total ($g_1 + g_2 = 6394$).

The species richness of quadrats was subject to considerable fluctuation. Species *turnover* (number of species disappearing after 1977 plus the number of species occurring in 2000 per quadrat) ranged from 12 to 48. Species *balance* (no. of species in 1977 minus no. of species in 2000 per quadrat) also varied greatly (from -17 to 19).

After examining the demography of each species, we find that there was a considerable rearrangement of frequencies behind the constant grand totals. Table 2 lists species, and their life forms, whose frequency decreased or increased by at least 10. It is striking that the "loser" species represent the perennial or geophytic life forms, whereas most "winners" belong to the annual (Th) or the woody (MM) category. The number of species disappearing from at least 5 quadrats after 1977 was 27, whereas the number of species appearing new in at least 5 quadrats in 2000 was 30.

Histograms

Figure 1a shows the life form histograms of the two study years. To satisfy the minimum cell-frequency requirements of the chi-square test, two categories on the left (Th + TH), and three categories on the right (N + M + MM) were combined, thus giving a total of 5 categories. The resulting chi-square value is 71.1 with $p < 0.0001$ for the null hypothesis, the homogeneity of the two distributions, being true. The main reason for rejecting the null hypothesis is that the proportion of annuals and biennials (Th + TH) and woody species (N + M + MM) increased on expense of the three intermediate life form categories. Species mostly responsible for this change are listed in Table 2.

The total numbers of occurrences of the eight seed mass categories for 1977 and 2000 are

Table 2. Species with the highest change of frequency during the study years in the Sas-hegy grassland vegetation. See text, for abbreviations of life-form classes.

Species	Life form	Frequency change
<i>Koeleria cristata</i> (L.) Pers.	H	-46
<i>Allium sphaerocephalon</i> L.	G	-44
<i>Melica ciliata</i> L.	H	-27
<i>Bromus pannonicus</i> Kumm. et Sendtn.	H	-24
<i>Galium glaucum</i> L.	H	-19
<i>Artemisia campestris</i> L.	Ch	-17
<i>Minuartia verna</i> (L.) Hiern.	H-Ch	-16
<i>Seseli osseum</i> Cr.	H	-15
<i>Hieracium auriculoides</i> agg.	H	-14
<i>Minuartia setacea</i> (Thuill.) Hay.	H-Ch	-13
<i>Allium moschatum</i> L.	H-G	-13
<i>Dianthus plumarius</i> L. ssp. <i>regis-stephani</i> (Rapcs.) Baksay	H	-13
<i>Trinia glauca</i> (L.) Dum.	H	-12
<i>Jurinea mollis</i> (L.) Rchb.	H	-12
<i>Saxifraga tridactylites</i> L.	Th	10
<i>Silene otites</i> (L.) Wib.	H	10
<i>Carex liparicarpus</i> Gaud.	G	10
<i>Arabis auriculata</i> Lam.	Th	13
<i>Viola arvensis</i> Murr.	Th	13
<i>Linum tenuifolium</i> L.	H	14
<i>Erophila verna</i> (L.) Chev.	Th	15
<i>Hornungia petraea</i> (L.) Rchb.	Th	17
<i>Fraxinus ornus</i> L.	MM	22
<i>Cerastium pumilum</i> Curt.	Th	23
<i>Holosteum umbellatum</i> L.	Th	24
<i>Stipa eriocalis</i> Borb.	H	24
<i>Melampyrum arvense</i> L.	Th	25
<i>Veronica praecox</i> All.	Th	27
<i>Reseda lutea</i> L.	TH-H	28
<i>Arenaria serpyllifolia</i> L.	Th	29

summarized in Figure 1b. Since the last category had a very low frequency, it was merged with the previous one before calculating the test statistics. The chi-square test for trend yields $\chi^2 = 0.446$ with $p = 0.504$, showing that there was no uniform temporal change in any direction along the seed mass gradient. However, the two seed mass spectra are significantly different ($\chi^2 = 46.41$, $p < 0.0001$). The reason is that species with high and low seed mass increased their frequency, whereas those with intermediate class membership decreased by 2000. This is not contradictory to the non-significant result of the trend test where changes in opposite directions were canceled out.

Data matrices

The Manhattan distance between the two data matrices ($d = 1940$) is equivalent to the total

species turnover for the 1977 and 2000 samples. This number is a sum of two components, the number of occurrences detected only in 1977 (958) and those found only in 2000 (982). Thus, on average each quadrat lost approximately 12 species ($958/80 = 11.975$) and gained approximately 12 other species ($982/80 = 12.275$) after 23 years.

The observation that the total number of occurrences practically did not change over time strongly supports our preference in favor of data randomization designs that rely on fixed grand totals. The summary of simulation results for the five null models is presented in Table 3. The actual value of d (1940) is obviously significantly *low* under any null hypothesis, yet there are considerable differences between the simulated distributions. The fully constrained case is the closest to the actual, closely followed by the probabilistic model and then by the model with fixed species totals. The remaining two,

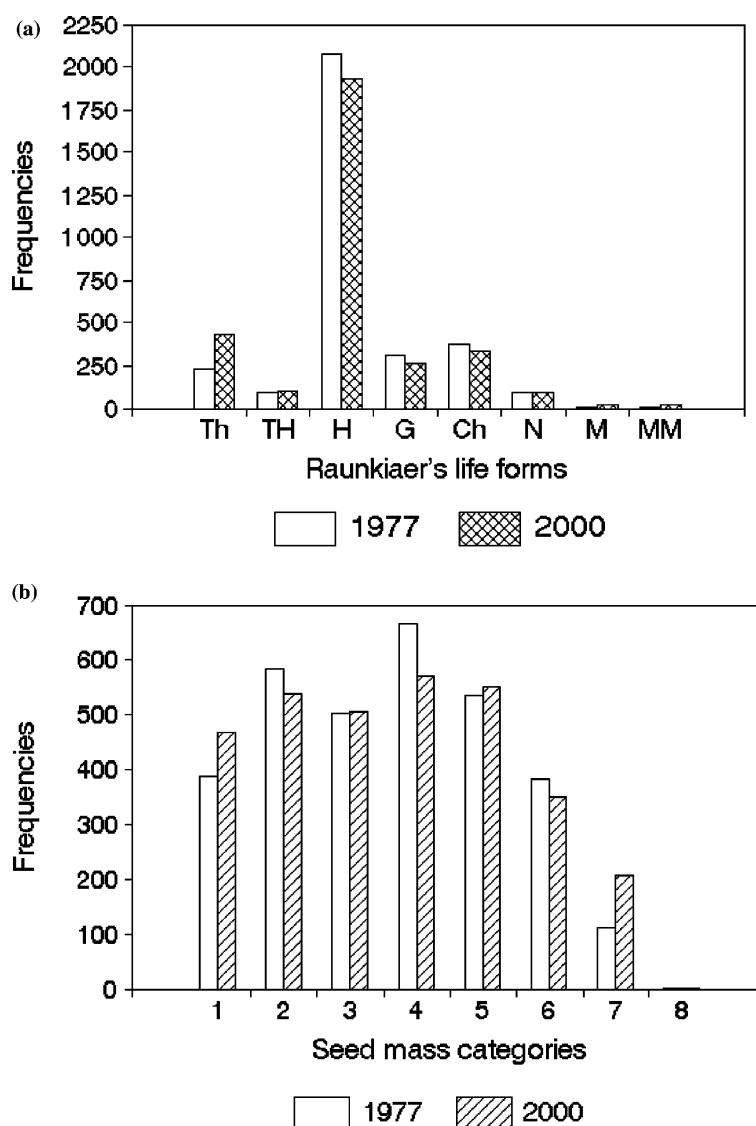


Figure 1. Life form (a) and seed mass (b) spectra in the grassland vegetation of Sas-hegy for the two study years. See text, for abbreviations of categories.

the unconstrained case and the model with fixed quadrat totals fall far apart, i.e., the random matrices obtained this way represent the highest departures from the actual data matrix.

Distance structures

The average Manhattan distance between quadrats is 31.98 for 1977 and 33.60 for 2000. These values, as rough measures of spatial floristic diversity indicate negligible differences between the

study years. Random permutations of the pooled distance matrix confirm this suggestion: the range of simulated averages is [30.95, 36.37] for either study year, i.e., the actual values are well inside this interval.

The mean distance between quadrats from different years (34.81) is only slightly higher than the within year averages. Yet, this value indicates a significant temporal difference based on the permutation test which provided a maximum value of 34.24 for 10 000 randomly rearranged matrices (i.e., $p < 0.0001$).

Table 3. Estimated parameters (minimum, maximum and mean) of the distribution of between-matrix Manhattan distances for five different randomization models.

	Unconstrained	Constant richness	Constant frequency	Full constraints	Probabilistic marginals
Minimum	4448	4424	2684	2578	2654
Maximum	4814	4744	2892	2794	2898
Mean	4655	4591	2786	2682	2776

The two distance matrices for quadrats have a matrix correlation of 0.651, which is much higher than the maximum (0.225) simulated by the Mantel test ($p \ll 0.0001$). The agreement between the two Manhattan distance matrices among species is even closer: the matrix correlation is 0.855 whereas the simulated maximum is 0.241.

Adjacency analysis

Five cases of spatiotemporal permanence, 107 cases of spatial permanence and 13 cases of spatiotemporal mutability were detected. Thirty-five quadrats had more than one nearest neighbors (tied minimum distances). Of these, 19 corresponded to spatial permanence because all neighbors came from the same year, 1 corresponded to spatiotemporal mutability, whereas 15 cases were associated with true ties, because the neighbors come from both years. The sum of these values, 35, was constant in the permutation tests and the associated quadrats are ignored here, because randomization modifies the positions of quadrat pairs (i.e., whether they come from the same or from different year). From the 10 000 simulations, the mean value of spatiotemporal permanence events was as low as 0.8 and only 70 outcomes proved to be at least 5 ($p = 0.007$). However, the number of cases of spatial permanence was always much lower in the simulations than 107, whereas the number of pairs showing spatiotemporal mutability always exceeded 15 greatly (with simulated means around 62 for both). This can only be explained by assuming that temporal compositional change was more substantial in affecting nearest neighbor relationships than spatial variation.

Ordinations and classifications

The PCoA ordinations of quadrats explain 45.2% (1977) and 40.1% (2000) of the total variance for

the first two axes, a fairly high percentage if we consider that the further eigenvalues decrease only very slowly. The two-dimensional PCoA ordinations for the two years have a dissimilarity of 0.39 (sum of squares, after normalization, uniform dilation and rotation in Procrustes analysis). With the second ordination randomly relabeled, the 5000 runs produced a mean sum of squares of 1.21, and a range of [1.03, 1.46], showing that the two ordinations are more similar than expected when one of them is randomized.

The joint PCoA ordination of all the 160 quadrats explains 25% and 10% of variation on the first two axes. Most trajectories connecting temporal pairs on the ordination plane are relatively short (Figure 2). There is no uniform alignment of the trajectories; suggesting that none of the axes can be interpreted as a temporal axis. Nevertheless, careful analysis of the diagram revealed some consistent behavior of trajectories when different parts of the ordination are examined separately. On the left side are the quadrats from N to NE exposition, with most arrows pointing to bottom right. Quadrats from gentle slopes in southern to southwestern exposition are located in the middle of the ordination and many of them show a move from bottom to the top. This group is not separated at all from the group of quadrats from steep southern slopes. Transitions into several directions are found. It is interesting that within this latter group temporal trajectories are the shortest and the least uniform in direction, demonstrating the relative stability of open grassland vegetation typical of steep southern dolomite slopes.

The complete linkage clustering of quadrats had a fairly high distance recovery, the cophenetic correlation being 0.751 for 1977 and 0.745 for 2000. The comparison of these dendrograms led to a result similar to the Procrustean comparison of ordinations. The matrix correlation for the two ultrametric matrices representing the dendrograms

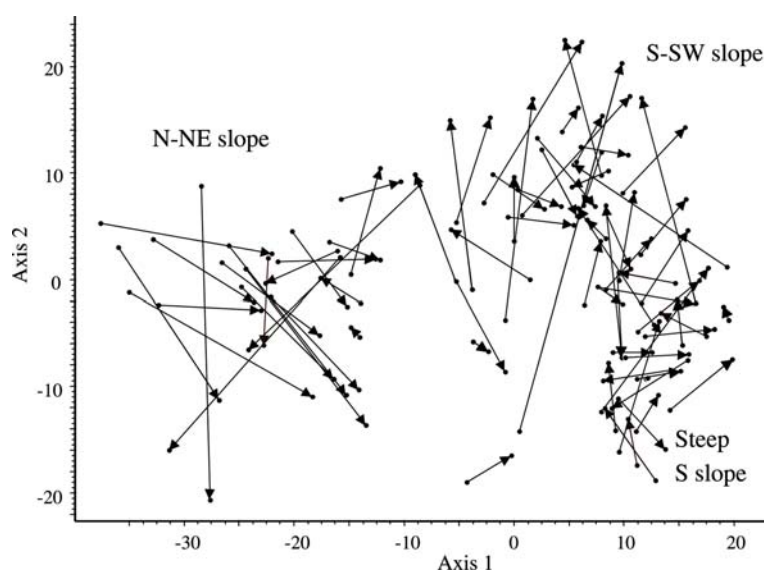


Figure 2. Principal coordinates ordination of Sas-hegy data with trajectories connecting temporal quadrat pairs. Arrows point towards the positions of quadrats sampled in 2000.

of sites is 0.578. The permutations yielded a zero mean and a range of $[-0.04, 0.05]$ for the matrix correlation, indicating that the two dendrograms are significantly similar. For comparing the dendrograms, we identified cut levels at about the same dissimilarity such that correspondence between clusters is the highest. It is enhanced by showing the geographical positions of the quadrats in the map of Figure 3. The two digits for each label indicate cluster memberships for the two years, respectively, and can be compared to see consistency of clustering and to identify rearrangements. The largest group of unchanged cluster membership has 27 quadrats, labeled “22”, and originates from the steep southern slopes. In the neighborhood of these quadrats, most of the labels are “11” and “12”, designating quadrats from gentle southern-southwestern slopes. Of the cluster-1 quadrats from 1977, 11 remained there in 2000, while 12 quadrats were moved into the group of southernmost exposition. Thus, the boundaries of the latter type apparently expanded by the year 2000. Classes 3 and 4 comprise quadrats from closed grasslands of northern-northeastern exposition. There is a notable rearrangement between these clusters, as shown by the opposite moves of sites labeled “43” and “34” and outlined in Figure 3. Almost all quadrats of the separate cluster 4 in 1977 became closer to the core of cluster 3, whereas the latter cluster was split as a

result of floristic changes. These results suggest that the open vegetation was less subject to successional changes.

For the combined classification of the 80 + 80 quadrats, the cophenetic correlation between the starting matrix and the hierarchical levels is 0.647. The decrease compared to the cophenetic correlations for the separate years is perhaps smaller than expected, if we consider that the number of objects in this classification is doubled. The classification into five groups is considered here, because of its partial congruence with the previous two analyses. Comparisons reveal that the most intensive changes in floristic composition occurred in the SW part of the study area, and within the closed grasslands of N-NW exposition. In the dendrograms for 1977 and 2000, 52.2% of quadrats have unchanged cluster memberships, whereas in the joint clustering approach 42 temporal pairs appear together. Results of 10,000 permutations of this partition suggest high significance of this value, because the simulated maximum was 36. That is, there are more temporal pairs remaining together after the 23 years than randomly expected. This indicates that in statistical sense the classification structure was not destroyed significantly during succession. Nevertheless, changes can be clearly outlined in the map, showing that these rearrangements have certain degree of directionality.

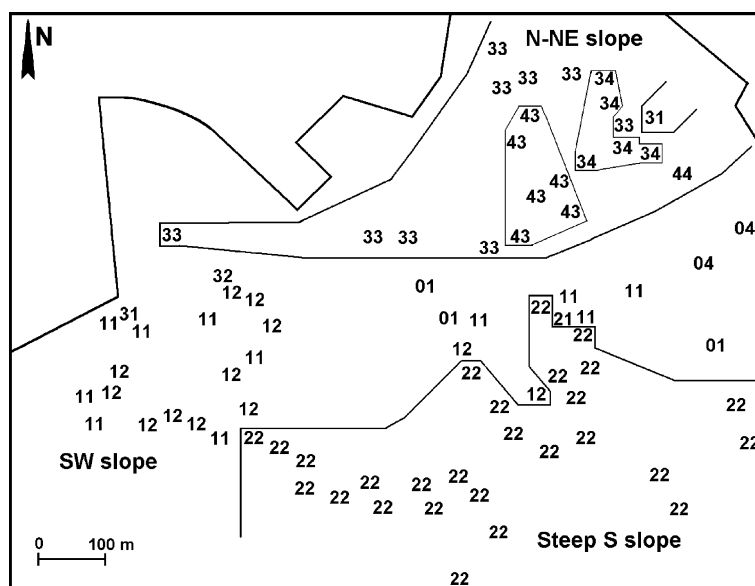


Figure 3. The geographic localization of 80 semi-permanent plots in Sas-hegy. For each quadrat position, the first digit refers to cluster membership in 1977, whereas the second digit corresponds to one of the clusters in 2000. Zero refers to small clusters ignored in the comparisons. The long boundary lines separate quadrats according to major groupings, discussed in text.

At this point, comparison of classifications with the results of adjacency analysis is in order. The 5 quadrats corresponding to spatiotemporal permanence fall into the same cluster as their nearest neighbors. Of the 107 quadrats showing the phenomenon of spatial permanence, 96 appear together with their nearest neighbors, and only 11 are separated into different clusters. Eleven out of the 13 quadrats showing spatiotemporal mutability are classified together with the respective nearest neighbors and only 3 are separated.

Discussion

Methodological aspects

We used two radically different approaches to test temporal change on the multivariate level: (1) for the data matrices, we proposed a novel method based on *direct randomization within the data body* with various constraints on marginal totals, and (2) *random relabeling* of objects in distance matrices, ordinations and classifications to yield a sample distribution of the test statistics.

Our method to detect significance in data structures is a new alternative to Procrustean superpositions advocated by Jackson (1995) and

Peres-Neto and Jackson (2001). We have shown that for at least two reasons, the use of Manhattan distance is more logical choice in comparing two alternative data matrices as a statistic to be tested. For the tests themselves, randomization algorithms widely used in assessing association of species in the presence/absence case (e.g., Sanderson et al. 1998; Stone and Roberts 1990; Gotelli 2000; Miklós and Podani 2004) provide flexibility, and offer several null hypotheses to be made in accordance with the expectations of the ecologist. Of the many possibilities, we tried five different designs that maintain the grand total which was surprisingly so similar in our example data sets. Incorporating other strategies without constraints on the grand total is also conceivable, of course, but we excluded them from the present study. The simulations provided results which can be assigned to two groups. Those implying uniform distribution of species (1 and 2) produced distributions falling the farthest from the actual data. These null models offered the most unrealistic situations, these produce most impossible species combinations, hence the highest deviation from the actual distance. In fact, equidistribution of species is practically never approximated in vegetation data and therefore these strategies are not recommended in general. The other three procedures are less restrictive,

although all indicate significant similarity of the two years. The null model implying uniform distribution for quadrat totals (3) is more realistic than algorithms 1–2, although quadrat totals are still far from the uniform. The probabilistic model with grand total adjustment (5) and the fully constrained (4) model are only a little closer to the actual value. The fully constrained case retains the most of the original matrix, yet the simulated variants fall very far from the original data. The tests show unambiguously that changing as many as 1940 scores in a matrix with a total of $80 \times 147 = 11760$ entries can be quite well within the range of changes that are expected by chance only.

Of course, the five randomization designs each could be followed by calculating distance matrices and then by comparing these matrices with the Mantel test. Furthermore, the distances may also be subjected to PCoA followed by Procrustes analysis of the resulting ordinations for a fixed number of axes, another possibility offered by Jackson (1995) and Peres-Neto and Jackson (2001). We felt, however, that these do not add anything new to the direct comparisons of data matrices and therefore we used a different strategy, i.e., random relabeling of objects to calculate the simulated statistics. This is the only way to reveal if the processes of compressing information into distances, ordinations and classifications produce something “new”, compared to the data level. One must be warned at this point that significant concordance detected in such an analysis has nothing to do with the starting compositional similarity. High concordance between two dendrograms, for example, is *not* the reflection of any agreement in species composition. We refer again to the two data matrices of Table 1, because they would produce identical distance matrices and therefore identical ordinations and classifications. At this level, we investigate only if the multidimensional *structure* of study objects has changed, and as we see from the artificial example, completely different subsets of species can produce identical summaries of data. This may appear trivial for some, yet we feel important to point this out explicitly.

The present investigation used a complex approach in which the high actual statistics may suggest a contradiction between different results. However, different null hypotheses require different explanation. In the χ^2 -based tests, the

null-hypothesis is that the data for the two years are statistically indistinguishable and differences are merely caused by sampling error and stochastic fluctuations. The extremely high actual statistics obtained indicate therefore significant differences between the two years. In permutation-based tests, the null-hypothesis is that the two observations are completely independent. The extreme values obtained now reflect that there are significant similarities, rather than differences, between the vegetation of two years.

Ecological aspects

Comparison of results of tests performed at different levels immediately shows that no single statement can be made whether the Sas-hegy vegetation has changed significantly after the 23 years. Tests involving the comparison of global vegetation pattern produced similar results. At the level of data matrices, none of the randomization designs revealed significant change, notwithstanding the high total species turnover between the two study years. Since the distance matrices and in turn the ordinations and classifications are all derived from the data, it is fairly acceptable that these do not exhibit significant temporal change either. The actual measures were always more extreme than those expected by chance. According to these findings, the rock grasslands of Sas-hegy appear to have a fair structural stability: there is no detectable temporal change of global pattern which suggests a high level of community organization.

Tests focusing on some minor or more detailed aspects of vegetation structure produced somewhat contrasting results, raising the possibility that behind global stability there are significant rearrangements overlooked by comparisons at the data/matrix/classification/ordination level. While total richness in the two years was similar, the disappearance of 8 species and appearance of 24 species seem to be substantial on ecological grounds. Of course, data from two years only are not sufficient to decide whether this net increase in number of species is a trend or fluctuation. At least for therophytes, great fluctuation was reported by Geißelbrecht-Taferner et al. (1997) within three years in similar fescue grassland communities from Austria. Considering the fact that most of these

species are accidental elements in the vegetation of Sas-hegy, one might be tempted to say that these changes are mere fluctuations. However, scrutiny of the seed mass and life form distributions and the list of Table 2 reveals contrasting ecological trends. Some phanerophytes, especially *Fraxinus ornus*, are more widespread in 2000, perhaps as an initialization towards forestation, a quite natural progressive change in grasslands (Begon et al. 1986; Dutoit et al. 2003; Bonet and Pausas 2004). However, this process is associated with the expansion of several annual and biennial species, whose appearance is generally considered as indication of colonization or degradation, rather than progression during later stages. For an andesite grassland community, Török et al. (1994) reported similar rearrangements in the life form distribution after more than 32 years – with a fair stability of species number – and attributed this at least partly to trampling effect. We cannot exclude the possibility that there was some trampling effect in our study area as well, despite all efforts to exclude visitors and dogs which could open up the habitat (some space) for these species. The comparison of seed mass spectra indicates similar trends. It has long been known that seed mass generally increases along with successional progression (Salisbury 1942; Leishman et al. 2000). The increase we observed at high seed mass categories corresponds with the raised frequency of phanerophytes, whereas the simultaneously observed increase in the lowest seed mass category agrees well with the annual/biennial expansion. These are changes that would have been undetected if one insisted on the multivariate approach to testing.

The proportional changes of hemicryptophytes are a different matter. Sas-hegy has a few perennial grasses declining during the study years, especially *Koeleria glauca*, *Bromus pannonicus* and *Melica ciliata*, whereas others, such as *Stipa eriocaulis*, show the opposite trend. It is well-known from the literature that perennial grasses can be subject to considerable fluctuation. For example, Watt (1971) reported that *Festuca ovina* exhibits a ten-year regeneration cycle: after a period of dominance the tussocks disintegrate and then the species is re-established. Although we do not have direct observations on such life cycles in Sas-hegy, we suppose that similar population dynamics are responsible at least partly for the fluctuations of perennials there.

Whereas Mantel tests did not show any significant overall temporal change of distances between quadrats, it is striking that the mean between-quadrat distance increased, slightly but significantly, between 1977 and 2000. When we focus our attention to a small subset of distances by searching for nearest neighbors in the pooled distance matrix, we again end up with a conclusion different from the overall comparisons of matrices. The number of quadrats that have their temporal pairs as nearest neighbors is not as high as might be expected from results that do not suggest significant temporal changes. Among the nearest neighbors, there are significantly more cases of spatial permanence and significantly less cases of spatiotemporal mutability than expected by chance. This shows that interesting details are masked by overall matrix comparisons: nearest neighbor relationships suffer a definite change between the study years. What is also striking is that adjacency analysis agrees quite well with the results of complete linkage clustering, a classificatory method emphasizing the opposite clustering criterion, i.e., the farthest neighbor relationships.

In many other studies of community change spanning 15–25 years, ordination trajectories reveal directional community change, and dimensions can be identified as temporal axes due to progressive succession (e.g., van der Maarel et al. 1985; Battaglia et al. 2002) or degradation (Török et al. 1994), but see Lepš et al. (1989) for the opposite in insect assemblages. In our case, none of the ordination dimensions can be identified as a temporal axis confirming the suggestion that vegetation change in Sas-hegy is more of a stochastic movement rather than change into a particular direction. This is true even if we can identify more trends when three groups of quadrats are examined separately. The remote positions of some temporal pairs in the joint ordination may be explained partly by the fact that the ordination on axes 1–2 cannot be an adequate representation of all pairwise distances. There is also a chance of sampling error caused by quadrats not positioned exactly in the same site in the two study years, despite our all efforts to exclude this possibility.

Comparisons of ordinations and classifications indicated no significant differences, that is, there is global temporal stability in the results. What is the explanation of the difference between the significance of global tests and those reflecting a

relatively minor aspect of vegetation structure? Multivariate data and their summaries in form of distance matrices, ordinations and hierarchical or non-hierarchical classifications are certainly more robust than derived variables or distributions. In order to reach significant changes in ordinations, for example, the vegetation studied should undergo considerable floristic and structural changes. We did not encounter such a strong rearrangement of community characteristics, and the changes in nearest neighbor relationships and life form and seed mass spectra, however significant, proved to be insufficient to cause substantial structural changes.

Acknowledgements

Thanks are due to D. M. Oosterheld and two anonymous referees for their constructive criticism of the manuscript. H. Punka is acknowledged for her assistance in seed mass measurements. We are grateful to the management of the Danube-Ipoly National Park for granting access to the study area in Sas-hegy. This work was supported financially by the Hungarian National Research Fund (OTKA grants T043732 and T037732).

References

- Altman D.G. 1991. *Practical Statistics for Medical Research*. Chapman and Hall, London.
- Austin M.P. 1977. Use of ordination and other multivariate descriptive methods to study succession. *Vegetatio* 35: 165–175.
- Bartha S., Rédei T., Szollát Gy., Bódis J. and Mucina L. 1998. Compositional diversity and fine-scale spatial patterns of dolomite grasslands on contrasting slopes. In: Csontos P. (ed.), *Botanical Exploration of Rock Grasslands. A Festschrift to B. Zólyomi*. Scientia, Budapest, pp. 159–182. (in Hungarian).
- Battaglia L.L., Michin P.R. and Pritchett D.W. 2002. Sixteen years of old-field succession and reestablishment of a bottomland hardwood forest in the Lower Mississippi Alluvial Valley. *Wetlands* 22: 1–17.
- Begon M., Harper J.L. and Townsend C.R. 1986. *Ecology: Individuals, Populations and Communities*. Blackwell, Oxford.
- Bonet A. and Pausas J.G. 2004. Species richness and cover along a 60-year chronosequence in old-fields of southeastern Spain. *Plant Ecol.* 174: 257–270.
- Clarke K.R. 1993. Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* 18: 117–143.
- Connor E.F. and Simberloff D. 1979. The assembly of species communities: chance or competition? *Ecology* 60: 1132–1140.
- Csontos P. 1998. The applicability of a seed ecological database (SEED) in botanical research. *Seed Sci. Res.* 8: 47–51.
- Csontos P. and Lőkös L. 1992. Védett edényes fajok térbeli eloszlás-vizsgálata a Budai hg. dolomitvidékén. – Szünbotanikai alapozás természetvédelmi területek felméréséhez. *Botanikai Közlemények* 79: 121–143.
- Csontos P., Tamás J. and Podani J. 2004. Slope aspect affects the seed mass spectrum of grassland vegetation. *Seed Sci. Res.* 14: 379–385.
- Debreczy Zs. 1966. Die xerothermen Rasen der Péter- und Tamás-Berge bei Balatonarács. *Annales Historico-naturales Musei Nationalis Hungarici* 58: 223–241.
- Debreczy Zs. 1987. Fluctuating dynamic equilibrium of photophil, xerophil rupicolous plant communities and scrub woods at the lower woodland limit. *Annales Historico-naturales Musei Nationalis Hungarici* 79: 89–112.
- Draskovits R. and Kovács-Láng E. 1968. Mikrokimamessungen in Kalkstein- und Dolomittfelsenrasen. *Annales Universitatis Scientiarum Budapestinensis, Sectio Biologica* 9–10: 115–129.
- Dutoit T., Buisson E., Roche P. and Alard D. 2003. Land use history and botanical changes in the calcareous hillsides of Upper-Normandy (north-western France): new implications for their conservation management. *Biol. Conserv.* 115: 1–19.
- Foster B.L. and Tilman D. 2000. Dynamic and static views of succession: Testing the descriptive power of the chronosequence approach. *Plant Ecol.* 146: 1–10.
- Geißelbrecht-Taferner L., Geißelbrecht J. and Mucina L. 1997. Fine-scale spatial population patterns and mobility of winter-annual herbs in a dry grassland. *J. Veg. Sci.* 8: 209–216.
- Gilpin M.E. and Diamond J.M. 1982. Factors contributing to non-randomness in species co-occurrences on islands. *Oecologia* 52: 75–84.
- Gotelli N. 2000. Null model analysis of species co-occurrence patterns. *Ecology* 81: 2606–2621.
- Gotelli N. and Entsminger G.L. 2001. Swap and fill algorithms in null model analysis: rethinking the knight's tour. *Oecologia* 129: 281–291.
- Gotelli N. and Graves G.R. 1996. *Null Models in Ecology*. Smithsonian Inst., Washington DC.
- Gower J.C. 1975. Generalized Procrustes analysis. *Psychometrika* 40: 33–51.
- InStat 1997. *GraphPad InStat Demo, Version 3.00 for Win 95/NT*. GraphPad Software Incl., San Diego.
- Jackson D.A. 1995. PROTEST: a PROcrustean randomization TEST of community environment concordance. *Ecoscience* 2: 297–303.
- Leishman M.R., Wright I.J., Moles A.T. and Westoby M. 2000. The evolutionary ecology of seed size. In: Fenner M. (ed.), *Seeds: The Ecology of Regeneration in Plant Communities*. CABI Publishing, Wallingford, pp. 31–57.
- Lepš J., Soldán T. and Landa V. 1989. Multivariate analysis of compositional changes in communities of Ephemeroptera (Insecta) in the Labe Basin, Czechoslovakia – A comparison of methods. *Coenoses* 4: 29–37.
- Maarel E. van der Boot R., van Dorp D. and Rijntjes J. 1985. Vegetation succession on the dunes near Oostvoorne, The Netherlands; a comparison of the vegetation in 1959 and 1980. *Vegetatio* 58: 137–187.
- Manly B.F.J. 1991. *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, London.

- Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H. and Teller E. 1953. Equations of state calculations by fast computing machines. *J. Chem. Phys.* 21: 1087–1091.
- Miklós I. and Podani J. 2004. Randomization of presence-absence matrices: comments and new algorithms. *Ecology* 85: 86–92. Supplement is available at Ecological Archives E085–001-S1.
- Peres-Neto P.R. and Jackson D.A. 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* 129: 169–178.
- Persson S. 1980. Succession in a south Swedish deciduous wood: a numerical approach. *Vegetatio* 43: 103–122.
- Philippi T.E., Dixon P.M. and Taylor B.E. 1998. Detecting trends in species composition. *Ecol. Appl.* 8: 300–308.
- Pickett S.T.A. 1989. Space-for-time substitution as an alternative to long-term studies. In: Likens G.E. (ed.), *Long-term Studies in Ecology: Approaches and Alternatives*. Springer-Verlag, New York, NY, pp. 110–135.
- Podani J. 1985. Syntaxonomic congruence in a small-scale vegetation survey. *Abstracta Botanica* 9: 99–128.
- Podani J. 1989. Comparison of classifications and ordinations of vegetation data. *Vegetatio* 83: 111–128.
- Podani J. 1992. Space series analysis of vegetation: processes reconsidered. *Abstracta Botanica*. 16: 25–29.
- Podani J. 1993. SYN-TAX version 5.0. User's Guide. Scientia, Budapest.
- Podani J. 1998. A complex numerical analysis of dolomite rock grasslands of the Sas-hegy Nature Reserve, Budapest, Hungary. In: Csontos P. (ed.), *Botanical Exploration of Rock Grasslands. A Festschrift to B. Zólyomi*. Scientia, Budapest, pp. 213–229. (in Hungarian).
- Podani J. 2000a. Simulation of random dendrograms and comparison tests: some comments. *J. Classif.* 17: 123–142.
- Podani J. 2000b. Introduction to the Exploration of Multivariate Biological Data. Backhuys, Leiden.
- Podani J. 2001. SYN-TAX 2000. Computer Programs for Data Analysis in Ecology and Systematics. User's Manual. Scientia, Budapest.
- Podani J., Csontos P. and Tamás J. 2000. Additive trees in the analysis of community data. *Community Ecol.* 1: 33–41.
- Podani J. and Dickinson T.A.D. 1984. Comparison of dendrograms: a multivariate approach. *Can. J. Bot.* 62: 2765–2778.
- Podani J. and Miklós I. 2002. Resemblance coefficients and the horseshoe effect in principal coordinates analysis. *Ecology* 83: 3331–3343.
- Salisbury E.J. 1942. *The Reproductive Capacity of Plants*. G. Bell and Sons, London.
- Sanders M.E., Dirkse G.M. and Slim P.A. 2004. Objectifying thematic, spatial and temporal aspects of vegetation mapping for monitoring. *Community Ecol.* 5: 81–91.
- Sanderson J.G., Moulton M.P. and Selfridge R.G. 1998. Null matrices and the analysis of species co-occurrences. *Oecologia* 116: 275–283.
- Schönemann P.H. and Carroll R.M. 1970. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika* 35: 245–256.
- Simon T. 2000. *A magyarországi edényes flóra határozója (Identification guide to the vascular plants of Hungary)*. Nemzeti Tankönyvkiadó, Budapest. (in Hungarian).
- Sneath P.H.A. and Sokal R.R. 1973. *Numerical Taxonomy*. Freeman, San Francisco.
- Stone L. and Roberts A. 1990. The checkerboard score and species distributions. *Oecologia* 85: 74–79.
- Török K., Horánszky A. and Kósa G. 1994. Long-term changes of species composition in an andesite grassland community of the Visegrád Mts., Hungary. *Abstracta Botanica*. 18: 13–27.
- Watt A.S. 1971. Factors controlling the floristic composition of some plant communities in Breckland. In: Duffey E. and Watt A.S. (eds), *The Scientific Management of Animal and Plant Communities for Conservation*. Blackwell Sci. Publ., Oxford, pp. 137–152.
- Zólyomi B. 1942. Die Mitteldonau Florenscheide und das Dolomitphänomen. *Botanikai Közlemények* 39: 209–231.
- Zólyomi B. 1958. Budapest és környékének természetes növénytakarója. (Natural vegetation of Budapest and its surroundings.). In: Pécsi M. (ed.), *Budapest Természeti Képe*. Akadémiai Kiadó, Budapest, pp. 508–642. (in Hungarian).