

# On randomization testing in vegetation science: multifactor comparisons of relevé groups

Pillar, Valério De Patta<sup>1</sup> & Orlóci, László<sup>2</sup>

<sup>1</sup>*Departamento de Botânica, Universidade Federal do Rio Grande do Sul, Campus Central, Porto Alegre, RS, Brazil 90046-900; Fax +55 51 2272295; E-mail pillar@vortex.ufrgs.br;*

<sup>2</sup>*Department of Plant Sciences, University of Western Ontario, London, Canada N6A 5B7*

**Abstract.** Hypothesis testing in phytocoenological applications is likely to be hindered when based on conventional statistical methods. The problem created by unrealistic assumptions can, however, be overcome by randomization. This paper discusses the general idea of randomization testing, describes a method and interprets its application in group comparisons. Two sets of variables are involved, the vegetation set on the basis of which the groups are compared and the environmental factors which delimit the groups under different analytical designs. Although simple partitioning of sum of squares is at the core of the test, the method has versatility of testing uni- or multifactor designs, which is novel in phytocoenological applications. The algorithm has been implemented in programs SYNCSA and MULTIV by V.P. Data from the Campos of southern Brazil are used for illustration.

**Keywords:** Campos; Community; Hypothesis; Interaction; Multivariate; Testing; Variance.

**Nomenclature:** Burkart (1969); Rosengurtt et al. (1970).

## Introduction

Standard statistical texts address hypothesis testing, most often in the context of classical 'Fisherian statistics' (Greig-Smith 1980). This approach requires the assumption that variables have a specific type of theoretical probability distribution, usually the Normal. The latter is now considered unduly confining (e.g. Bradley 1968:6; Edgington 1987:13; Potvin & Roff 1993; Orlóci 1993). As a matter of fact, Edgington (1987:5) pointed out that even if a normal distribution exists in the population, it is still an unsuitable model for hypothesis testing since random sampling that could capture the 'normal' structure is rarely implementable.

Randomization testing, often referred as 'permutation methods' is known since long (e.g. Pitman 1937; Fisher 1951:43; Kempthorne 1952: 128), but the technique was not practical then. This restricted the formulation of the problem and narrowed the definition of an 'acceptable' sampling environment (Orlóci 1993).

With the availability of fast computers, the advantages of the axiomatic approach largely dissipated and randomization methods became a practical proposition. Such methods were described in Bradley (1968) and Edgington (1969, 1987), but their emphasis is on the analysis of experimental univariate data. There are precedents applicable to multivariate samples. For example, Mantel (1967) presented a randomization test for comparing distance matrices (see also Smouse et al. 1986; Oden & Sokal 1992). Manly (1991), Crowley (1992) and Potvin & Roff (1993) gave reviews, including also other Monte Carlo methods, such as the bootstrapping method of Diaconis & Efron (1983) for estimation of confidence intervals.

Crowley (1992) referred to the class of Monte Carlo and randomization methods collectively as resampling methods. Particularly relevant to phytocoenology are the resampling methods used in population biology (e.g. Orlóci & Beshir 1976), community ecology (e.g. Orlóci & Kenkel 1985:80; Orlóci et al. 1986; Legendre & Fortin 1989; Podani 1991; Cornelius & Reynolds 1991; Pillar & Orlóci 1993a:69, ter Braak & Wiertz 1994) and ecosystem analysis (Carpenter et al. 1989).

Phytocoenological data analysis is essentially comparative. The relevés are observation vectors including vegetation and environmental variables as part of the total record from a site. Whether the relevés are taken individually or in groups, randomization tests can supply the probabilities needed for a straightforward significance test of the resemblance values.

In this paper we deal with the theory to an extent necessary for the practitioner in phytocoenology, describe an algorithm, comment on others, and present results from Campos vegetation. We see randomization testing in comparisons of relevé groups analogous to the tests in experimental designs involving multivariate analysis of variance. We rely on simple way of partitioning and testing variation in a manner which has advantage over others found in Edgington (1987), Manly (1991) and ter Braak & Wiertz (1994).

## Method

### The Null Hypothesis

Hypothesis testing involves probabilistic assessments of how common an observed value of the test criterion would be if a chosen Null Hypothesis,  $H_0$ , were in fact true. In vegetation data analysis the test criterion may apply to two or more groups of relevés and  $H_0$  may stipulate random arrangements of relevés among groups. The test criterion may compare the groups using the vegetation or the environmental data. For  $H_0$  to be relevant some alternative hypothesis or hypotheses must be in view where these are expected to produce distributions for the test criterion that are different from that obtained when the null hypothesis is true.

In randomization, a data set is visualized as one of many possible permutations of data elements among the available locations in the data matrix. The collection of all possible permutations represents a *reference set* (Hope 1968; Edgington 1987:305). This set can be generated by systematically shifting data elements between locations in ways that agree with  $H_0$ . Edgington (1987:43) described this as *systematic data permutation*, under which each permutation is just as likely as the observed data when  $H_0$  is true. But the reference set may be too large for direct computation. The solution to overcome this problem is found in the use of a limited but still large number of random permutations generated iteratively by an appropriate algorithm (Dwass 1957; Hope 1968; Edgington 1987:43). The observed data set is counted as one iteration.

The test involves the observed value  $\lambda$  of the test criterion and an associated distribution under the  $H_0$ . This distribution allows us to define the probability  $P$  that any randomly chosen value  $\lambda^\circ$  from among those that materialize in the permuted set will be at least as extreme as  $\lambda$ . When the test is concerned only with values that are at least as large as  $\lambda$ :

$$P(\lambda^\circ \geq \lambda) = \frac{\text{number of iterations in which } \lambda^\circ \geq \lambda}{\text{total number of iterations}} \quad (1)$$

The minimum of this is 1/total number of iterations, that is, in at least one iteration  $\lambda^\circ$  is the observed value.  $H_0$  is accepted if  $P$  is larger than some threshold value  $\alpha$ . Hope (1968) shows that, if the number of iterations is large, a random sample of the reference set will give reliable results, that is,  $P$ -values close to the exact ones that would be obtained in complete systematic data permutation. Crowley (1992) suggested a minimum of 1000 and 5000 iterations for tests at the 5% and 1% significance levels (cf. Edgington 1987:43; Manly 1991:15).

### Test criterion

The test we outline is applicable to data from sampling units described by any number of variables. In fact it is irrelevant what type and how many variables are involved, provided that an adequate dissimilarity measure can be found to gauge the level of resemblance in pairs of sampling units (relevés). In a sum of squares based test, the test criterion is the sum of squared dissimilarity between groups:

$$Q_b = Q_t - Q_w \quad (2)$$

In this,

$$Q_t = \frac{1}{n} \sum_{h=1}^{n-1} \sum_{i=h+1}^n d_{hi}^2 \quad (3)$$

which is the total sum of squares involving  $n(n-1)/2$  pair-wise squared dissimilarities ( $d_{hi}^2$ ) between  $n$  relevés. The  $Q_w$  term is given by

$$Q_w = \sum_{c=1}^k Q_{wc} \quad (4)$$

This is the sum of squares within the  $k$  groups. The terms in  $Q_w$  accord with

$$Q_{wc} = \frac{1}{n_c} \sum_{h=1}^{n-1} \sum_{i=h+1}^n d_{hi}^2 \delta(h, i, c) \quad (5)$$

which is the sum of squares within group  $c$  with size  $n_c$ . The indicator variable  $\delta(h, i, c)$  is one if relevés  $h$  and  $i$  belong to group  $c$  or zero if otherwise. The additivity of the terms is valid if the dissimilarity matrix has Euclidean metric properties; otherwise the algebra in the computation of  $Q_b$  and further partitioning will be unacceptable. Orlóci (1978) and Gower & Legendre (1986) give guidelines about measures with these properties. Similar partitions of sum of squares were used earlier by Ward (1963), Edwards & Cavalli-Sforza (1965) and Orlóci (1967, 1978:239), but for different purposes.

Computing sum of squares from the dissimilarity matrix has an advantage: it does not require the computation of group centroids from coordinates and deviations from centroids as used by i.a. Edgington (1987:190) and Manly (1991). Also, it is not constrained by the type of variables that describe the vegetation units, provided that the dissimilarity measure meets the requirements, a property especially useful in mixed data types.

Designs with two or more factors require the computation of interactions. Relevé groups may be defined by the states of the factors considered in isolation or jointly.

For instance, if factors  $X$  and  $Y$  have two states each, there will be four relevé groups based on the joint states. The joint classification defines a sum of squares as already explained, which can be partitioned as:

$$Q_{b|X+Y} = Q_{b|X} + Q_{b|Y} + Q_{b|XY} \quad (6)$$

$Q_{b|X}$  and  $Q_{b|Y}$  are the sum of squares between groups specific to factors  $X$  and  $Y$  considered in isolation.  $Q_{b|XY}$  is the interaction term (taken as a difference). Fig. 1 illustrates the computations. Similar manipulations apply when several factors are considered. A different  $Q_b$  is computed for the joint groups and for the groups according to each factor taken individually, pair-wise, in triplets, or on higher levels, but then manipulations can become very complicated. It must be noted that the computation of an interaction is only feasible when the data set contains records for all possible combinations of factor states.

The test may involve multiple contrasts in the same sense as in Scheffé (1953). A contrast is defined by positive and negative contrast coefficients indicating the groups of vegetation units to be compared. For any contrast  $j$  one computes  $Q_{ij}$ , the total sum of squares,  $Q_{wj+}$ , the sum of squares within groups with positive

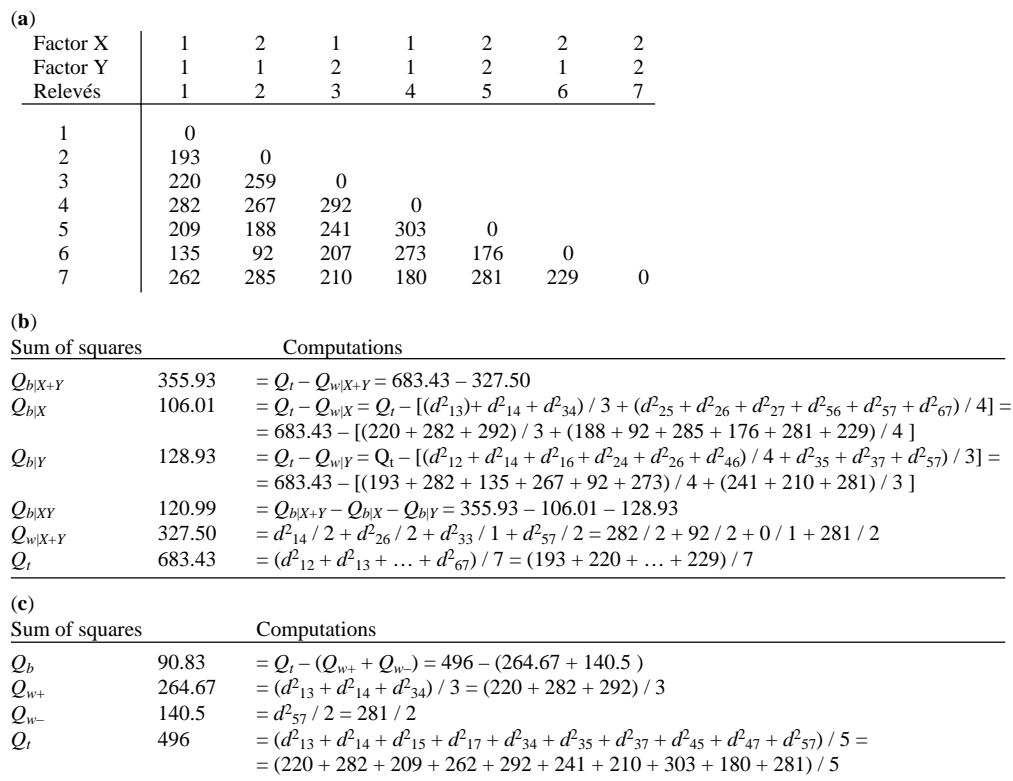
contrast coefficients, and  $Q_{wj-}$ , the sum of squares within groups with negative coefficients. The between groups sum of squares for contrast  $j$  is then  $Q_{bj} = Q_{ij} - (Q_{wj+} + Q_{wj-})$ . An illustration is given in Fig. 1c. For  $u = k - 1$  independent contrasts,

$$Q_b = \sum_{j=1}^u Q_{bj}$$

Provided the coefficients of a contrast are so assigned that they sum zero, two contrasts are independent if the inner product of the vectors containing the coefficients is zero.

*The reference set*

With a set of relevés, partitioned in groups according to external criteria, such as some environmental variables or experimental treatments, a relevant question is whether the observed compositional difference between the groups could have arisen by chance alone. The pertinent  $H_0$  must therefore stipulate that the observed classification of the relevés is a *random partition*. Where the groups refer to treatments, as in an experiment,  $H_0$



**Fig. 1.** Example with artificial data to illustrate the computations of sum of squares in a multifactor design. In (a) the matrix contains the squared Euclidean distances of seven relevés. The relevés are grouped by the states of factors X and Y. The sum of squares accord to equations in the main text. Table (b) gives computational results. Table (c) computes a contrast with coefficients 1 1 0 -2 for the groups 11 12 21 22 defined by the joint states of the two factors. This contrast involves relevés 1, 3, 4, 5, 7.

will stipulate that no differential treatment effect is evident in the measured compositional structure.

When the factors are not nested, which is the case in complete randomized designs with any number of factors, the reference set of the test is defined by unrestricted permutation of the observation vectors, our relevés, among the group memberships. This is equivalent to permuting the group memberships among the relevés. If a data set has three relevés partitioned in two groups, where the first two belong to group A and the last one to group B, there are six permuted sets:

Relevés:	123	123	123	123	123	123
Group memberships:	AAB	ABA	AAB	ABA	BAA	BAA

It is seen that in this example there are two identical copies of each of three partitions (AAB, ABA, BAA), but this does not affect the test. The probability of finding a specific partition by chance is still 1/3, irrespective of repetitions in the reference set. It is therefore implied that the random partitions algorithm can generate any element in the reference set simply by randomly allocating  $n$  relevés among  $n$  group memberships, without being concerned about repetitions. Actually, for  $n$  relevés partitioned in  $k$  groups, with the  $i$ th group receiving  $n_i$  relevés, the complete reference set will contain  $n!$  permutations, each one repeated

$$\prod_{i=1}^k (n_i!) \quad \text{times.}$$

In multifactor analytical designs the group memberships among which relevés are exchanged in randomization are defined by joint group memberships based on the states of the factors. For example, let memberships accord to:

Relevés:	1	2	3	4	5	6	7	8
States of factor X in relevés:	1	2	1	1	2	2	2	2
States of factor Y in relevés:	1	1	1	2	1	1	2	2
Then the joint memberships are coded as:	11	21	11	12	21	21	22	22

A possible random permutation of observation vectors among the groups will be:

Relevés:	2	3	8	5	7	4	6	1
----------	---	---	---	---	---	---	---	---

For the sake of completeness we mention also that in block designs and other nested designs the reference set is defined by restricting random allocations to the groups within the blocks (Edgington 1987:148; Manly 1991:67). For instance, eight relevés are in three groups (FGH) of factor N nested within three blocks (ABC) defined by factor M (block C is incomplete):

Relevés:	123	456	78
Blocks (factor M):	AAA	BBB	CC
Factor N:	GFH	FHG	HG

An element of the reference set is generated by randomly allocating  $n_h$  relevés among the groups of factor N within each block h of factor M. A possible permutation in the example above is:

Relevés:	321	645	87
Blocks (factor M):	AAA	BBB	CC
Factor N:	GFH	FHG	HG

It should be noted that this is adequate for testing the effect of any number of factors that are below the level of blocks without further nesting. For testing the effect of factors that are defining blocks there must, of course, be replicated blocks. If the number of units in each block is the same, the blocks can be permuted as a whole. If not, the units within blocks will have to be averaged and then permuted, which will require a separate test.

#### *The modus operandi abbreviated*

The algorithm takes as input the group memberships and the dissimilarity matrix of the relevés. At each iteration the randomization algorithm generates a random element of the reference set, rearranges the dissimilarity matrix accordingly, computes the  $Q_{bs}$  that are relevant to the analytical design, and compares the  $Q_{bs}$  with the corresponding ones actually observed, as already explained in general terms. After a large number of iterations, including the observed data set, the probability of an at least as extreme  $Q_b$  as the observed is the proportion  $P(Q_b^o \geq Q_b)$ . Fig. 2 illustrates the process. The results are presented and interpreted in a similar way as in a conventional analysis of variance.

#### *Example*

We use data from relevés taken in grassland communities, in an area covering 30 ha in the vicinity of Porto Alegre, in the Campos vegetation formation (Cabrera 1971). The site is grazed and floristically rich (165 species). The soils are yellow-red latosols on convex slopes and hydromorphic on low-lying terrain. The sample contains 60 quadrats, 0.25 m<sup>2</sup> each, along elevation gradients in the rolling landscape. The data set contains estimates of cover-abundance – van der Maarel's (1979) 1-9 ordinal transform scale – for 60 species that were present in at least 50% of the quadrats in at least one of the vegetation types. The 60-quadrat and 60-species matrix configuration is coincidental. Environmental data include records of landscape position, grazing condition, soil texture, soil pH and soil macro and micro nutrients. The complete data set is in Pillar (1988).

Relevé groups are defined a posteriori on the basis of environmental factors. We use landscape position to classify the quadrats into four groups (flat top, convex

(Iteration 1, observed data)

Factor X	1	2	1	1	2	2	2
Factor Y	1	1	2	1	2	1	2
Relevés	1	2	3	4	5	6	7
1	0						
2	193	0					
3	220	259	0				
4	282	267	292	0			
5	209	188	241	303	0		
6	135	92	207	273	176	0	
7	262	285	210	180	281	229	0

(Iteration 2)

Factor X	1	2	1	1	2	2	2
Factor Y	1	1	2	1	2	1	2
Relevés	1	5	4	2	6	7	3
1	0						
5	209	0					
4	282	303	0				
2	193	188	267	0			
6	135	176	273	92	0		
7	262	281	180	285	229	0	
3	220	241	292	259	207	210	0

... repeat the steps a large number of times.

In a run with 1000 iterations the results were:

Observed sum of squares	Proportion of iterations in which $Q_b^o \geq Q_b$
$Q_{b X+Y}$ 355.93	0.322
$Q_{b X}$ 106.01	0.580
$Q_{b Y}$ 128.93	0.273
$Q_{b XY}$ 120.99	0.378
$Q_{w X+Y}$ 327.50	not tested
$Q_t$ 683.43	not tested

slope, concave slope, lowland) and grazing intensity to classify the same quadrats into two groups (low, high). Exploratory ordination of relevés (Fig. 3) reveals trends associated with these factors. The question asked is if compositional differences are independent of landscape position and grazing intensity. The randomization is performed under random partition hypothesis. The algorithm computes sum of squares and probabilities based on the same matrix of chord distances that produced the ordination. We present these in Table 1. The results indicate a significant difference of vegetation composition between relevé groups defined by landscape and grazing, both jointly and independently. The landscape  $\times$  grazing interaction term, however, and one of the orthogonal contrasts (flat top versus convex slope) are not significant.

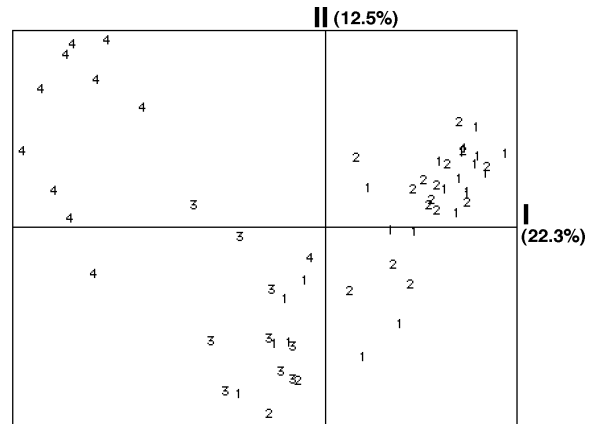
### Discussion

We have discussed randomization testing, described a method for group comparisons and have given an example from phytocoenology. Our method supports

Sum of squares	Is $\geq Q_b$ observed?
$Q_{b X+Y}$ 355.93	yes
$Q_{b X}$ 106.01	yes
$Q_{b Y}$ 128.93	yes
$Q_{b XY}$ 120.99	yes
$Q_{w X+Y}$ 327.50	not tested
$Q_t$ 683.43	not tested

Sum of squares	Is $\geq Q_b$ observed?
$Q_{b X+Y}$ 342.93	no
$Q_{b X}$ 100.10	no
$Q_{b Y}$ 71.59	no
$Q_{b XY}$ 171.24	yes
$Q_{w X+Y}$ 340.50	not tested
$Q_t$ 683.43	not tested

**Fig. 2.** Example with artificial data in a multifactor design to illustrate the randomization process under random partitions. In iteration 1 the squared distance matrix is arranged as observed. In the successive iterations the relevés are permuted at random and the dissimilarities rearranged accordingly, but the group memberships remain at the same positions in the table. At each iteration the test criterion is recomputed.



**Fig. 3.** Ordination diagram of Campos grassland relevés (Pillar 1988). The ordination is based on PCOA of a matrix of Euclidean chord distances. Numbers in the scattergram identify landscape classes: 1 = flat top; 2 = convex slope; 3 = concave slope; 4 = lowland. Species most correlated with axis 1 are *Paspalum notatum* and *Facelis retusa*, increasing from left to right, and *Eleocharis glauco-virens*, increasing from right to left. On axis 2, *Andropogon lateralis*, *Eryngium horridum* and *Vernonia nudiflora* decrease from bottom to top. Less grazed relevés are in the lower part of the diagram, more grazed ones are in the upper part.

**Table 1.** Randomization test of the independence of floristic composition from landscape and grazing intensity in Campos grassland. See data description in the main text. Relevés (60) are assigned to groups defined by the combination of four landscape categories and two grazing intensity levels. Probabilities ( $P(Q_b^\circ \geq Q_b)$ ), generated in 1000 iterations, indicate how extreme is the observed sum of squares under the Null Hypothesis. The analysis is akin to a conventional analysis of variance, differing only in the way probabilities are obtained. The sum of squares are computed from chord distances of relevés based on 60 species. The probability ( $P(Q_b^\circ \geq Q_b) = 0.001$ ) found for the sum of squares between groups defined by landscape and grazing indicates the rejection ( $\alpha \leq 0.05$ ) of the Null Hypothesis. Contrast coefficients specify which landscape classes (flat top, convex slope, concave slope, lowland) are compared. The contrasts are orthogonal, thus their sum of squares add to the sum of squares between groups by landscape ( $6.270 + 2.419 + 0.8326 = 9.521$ ). The probabilities for contrasts show significant compositional differences ( $\alpha \leq 0.05$ ) of lowland relevés versus the other groups taken together (1st contrast), and of concave slope relevés versus flat top plus convex slope relevés (2nd contrast), but not of flat top versus convex slope relevés (3rd contrast). Variation related to grazing intensity is also significant ( $\alpha \leq 0.05$ ), but the interaction of landscape and grazing is not ( $P(Q_b^\circ \geq Q_b) = 0.884$ ). The between groups sum of squares terms are additive; the  $Q_b$  for landscape plus the  $Q_b$  for grazing and the interaction add to the  $Q_b$  for grazing and landscape taken jointly ( $9.521 + 3.271 + 1.506 = 14.30$ ). The total sum of squares is the sum of the within plus the between groups by grazing and landscape taken jointly ( $24.18 + 14.30 = 38.48$ ). The vegetation variation explained by the factors is 37% ( $100 \times 14.30 / 38.48$ ) of the total.

Source of variation	Sum of squares ( $Q_b$ )	$(P(Q_b^\circ \geq Q_b))$
Between groups by landscape/grazing	14.30	0.001
Between groups by landscape	9.521	0.001
Contrasts:		
1 1 1 -3	6.270	0.001
-1 -1 2 0	2.419	0.002
1 -1 0 0	0.8326	0.181
Between groups by grazing	3.271	0.001
Interaction landscape $\times$ grazing	1.506	0.884
Within groups by landscape/grazing	24.18	
Total	38.48	

hypothesis testing in a specific case that could not have been done as well and as relevantly without randomization.

The method involves permutations of the relevés among the group memberships. It is clear that such a permutation is equivalent to permuting the relevés among the sites, but keeping the group memberships of the sites unchanged. It is also true that the test performed has analog in Mantel's (1967) test of matrix correspondence when the resemblance matrices are defined by two sets of variables. Legendre & Fortin (1989) tested vegetation-environment associations on this basis. One of the sets in their case is from the vegetation and the other from the environment or from geographical coordinates. A suitable test criterion in the Mantel test is structural

congruence  $\rho(\mathbf{D}; \Delta)$  defined as a matrix correlation, similarly as in Sokal & Rohlf (1962) and Orlóci & Pillar (1989). Given  $n$  pairs of vegetation and environmental relevés,  $\mathbf{D}$  defines vegetation structure, as an  $n \times n$  matrix of dissimilarities, and  $\Delta$  defines the underlying environmental structure, also as an  $n \times n$  matrix of dissimilarities. In agreement with the Null Hypothesis of no structural congruence, the reference set contains  $n!$  permutations of the  $n$  vegetation relevés among the  $n$  sites. The permutation of environmental relevés is unnecessary. Therefore, the reference set is equivalent to the one defined under the random partition hypothesis.

While the reference sets are equivalent in Mantel's test and in the random partition test, the test criteria differ. While in the latter  $Q_{bs}$  measure dispersion be-

**Table 2.** Randomization test of the independence of floristic composition from landscape in Campos grassland. Data contain 16 relevés, which are pooled relevés taken from the ones used in Table 1 that are at the same landscape position and in the same elevation gradient. Pooling the relevés aims at isolating landscape from a possible confounding effect of spatial proximity. Some probabilities  $P(Q_b^\circ \geq Q_b)$  differ from Table 1 for the same comparisons, but not enough to change the conclusions there stated about the association of landscape and vegetation. Probabilities were generated by 1000 iterations.

Source of variation	Sum of squares ( $Q_b$ )	$(P(Q_b^\circ \geq Q_b))$
Between groups by landscape	4.458	0.001
Contrasts:		
1 1 1 -3	2.886	0.002
-1 -1 2 0	1.324	0.014
1 -1 0 0	0.2484	0.931
Within groups by landscape	3.304	
Total	7.762	

tween groups, in the former  $\rho(\mathbf{D}; \Delta)$  measures correlation. Interestingly, the factors that define the groups in the computation of the  $Q_b$  could define  $\Delta$ . The advantage would be that the test would not be limited to categorical factors. The case would then be handled as has been done in Smouse et al. (1986). Along similar lines are various applications of canonical analysis in Gittins (1979, 1985) and ter Braak (1986, 1994), albeit not in the context of randomization tests as in Borcard et al. (1992) and ter Braak & Wiertz (1994).

The approach based on  $Q_b$  has advantages. Unlike  $\rho(\mathbf{D}; \Delta)$ , it is partitionable into components. One of these, the interaction term, adds an extra dimension to the partitioning in that it allows the complete evaluation of multifactor designs. This puts the analysis on a par with the factorial analysis of variance. Furthermore,  $Q_b$  allows contrasts to be defined and tested and yet the algorithm is not complicated.

If the resemblance matrix does not have Euclidean metric properties, other test criteria could still be used in some instances, such as an average dissimilarity between groups (Pillar & Orlóci 1993a:108) or our  $Q_w$  taken as a measure of within group dispersion, but these are not partitionable.

Some precautions are needed when tests involve survey data. First, the nature of the data excludes drawing valid conclusions about causality. Second, even though significance is indicated, other unknown factors may account for a large portion of total variation. Upon examination of the sum of squares in Table 1, we see that vegetation variation explained by landscape and grazing combined is only 37%. This is indicating that in addition to landscape and grazing, there must be other factors influencing vegetation variation. Related to this we refer to Quinn & Dunham (1983) for further comments on the hypothesis testing controversy.

Another important point is that if relevés from sites that are closer in space and in the same group tend to be less dissimilar, there will be a confounding effect of environmental factors and biological processes, such as the ones related to plant dispersion. In this case spatial autocorrelation may be significant and the probabilities will likely be underestimated, which may cause misleading conclusions on effects of environmental factors. Smouse et al. (1986), Legendre & Fortin (1989), Oden & Sokal (1992), Sokal et al. (1993), Legendre (1993) address the problem of spatial autocorrelation by means of a partial Mantel test and canonical correlation. We believe that the confounding effect of spatial autocorrelation can be alleviated by averaging the relevés in the same group when they are at close spatial proximity. It can be argued that autocorrelation affects the results in the Campos example in the sense that landscape and spatial proximity could be confounded.

As a precaution, in a second run we pooled relevés at similar landscape positions and in the same elevation gradient, which resulted in a new data set. This does not change the conclusions, which suggests the lack of confounding effects (Table 2). This solution would not be useful, however, if each group of relevés were representing a geographically contiguous area. Legendre et al. (1990) and Sokal et al. (1993) offer a method based on a contiguity-constrained permutation of units among groups, but the method requires evenly spaced relevés and a regular number of edges per contiguous area. A restricted randomization of this type could be implemented using the same partitionable test criterion we use.

It should be noted that perceptions of the vegetation phenomena tend to be context sensitive. Conclusions based on phytocoenological analysis are always dependent on a number of decisions about vegetation sampling: scale, size, unit shape and size (Kenkel et al. 1989). Also, the taxonomy delimiting populations matters (Pillar & Orlóci 1993b) and the choice of data transformation and resemblance measure is important. Therefore, the probabilities generated in randomization will also be sensitive to this choices.

**Acknowledgements.** Financial support from CNPq (Brazil) to V. Pillar and from NSERC (Canada) to L. Orlóci is gratefully acknowledged. Thanks are extended to M. E. Biondini and anonymous reviewers for helpful suggestions. For specifications and availability of the SYNCSA and MULTIV packages, please contact V. Pillar.

## References

- Borcard, D., Legendre, P. & Drapeau, P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73:1045-1055.
- Bradley, J.V. 1968. *Distribution-free statistical tests*. Prentice-Hall, Englewood Cliffs, NJ.
- Burkart, A. (ed.) 1969. *Flora ilustrada de Entre Rios (Argentina)*. INTA, Buenos Aires.
- Cabrera, A.L. 1971. *Fitogeografía de la República Argentina*. Sociedad Argentina de Botánica, Buenos Aires.
- Carpenter, S.R., Frost, T.M., Heisey, D. & Kratz, T.K. 1989. Randomized intervention analysis and the interpretation of whole-ecosystem experiments. *Ecology* 70: 1142-1152.
- Cornelius, J.M. & Reynolds, J.F. 1991. On determining the statistical significance of discontinuities within ordered ecological data. *Ecology* 72: 2057-2070.
- Crowley, P.H. 1992. Resampling methods for computation-intensive data analysis in ecology and evolution. *Annu. Rev. Ecol. Syst.* 23: 405-447.
- Diaconis, P. & Efron, B. 1983. Computer-intensive methods in statistics. *Sci. Am.* 248: 96-108.
- Dwass, M. 1957. Modified randomization tests for nonpara-

- metric hypotheses. *Ann. Math. Stat.* 28: 181-187.
- Edgington, E.S. 1969. *Statistical Inference: The Distribution-Free Approach*. McGraw-Hill, New York, NY.
- Edgington, E.S. 1987. *Randomization tests*. Marcel Dekker, New York, NY.
- Edwards, A.W.F. & Cavalli-Sforza, L.L. 1965. A method for cluster analysis. *Biometrics* 2: 362-375.
- Fisher, R.A. 1951. *The design of experiments*. 6th ed. Oliver and Boyd, Edinburgh.
- Gittins, R. 1979. Ecological applications of canonical analysis. In: Orlóci, L., Rao, C.R. & W. Stiteler, W.M. (eds.) *Multivariate methods in ecological work*, pp. 309-535. International Co-operative Publishing House, Fairland, MD.
- Gittins, R. 1985. *Canonical analysis; a review with applications in ecology*. Springer-Verlag, Berlin.
- Gower, J.C. & Legendre, P. 1986. Metric and euclidean properties of dissimilarity coefficients. *J. Clas.* 3: 5-48.
- Greig-Smith, P. 1980. The development of numerical classification and ordination. *Vegetatio* 42: 1-9.
- Hope, A.C.A. 1968. A simplified Monte Carlo significance test procedure. *J. R. Stat. Soc.* 30: 582-598.
- Kempthorne, O. 1952. *The design and analysis of experiments*. J. Wiley, New York, NY.
- Kenkel, N.C., Juhász-Nagy, P. & Podani, J. 1989. On sampling procedures in population and community ecology. *Vegetatio* 83: 195-207.
- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74: 1659-1673.
- Legendre, P. & Fortin, M.-J. 1989. Spatial pattern and ecological analysis. *Vegetatio* 80: 107-138.
- Legendre, P., Oden, N.L., Sokal, R.R., Vaudor, A. & Kim, J. 1990. Approximate analysis of variance of spatially autocorrelated regional data. *J. Class.* 7: 53-75.
- Manly, B.F.J. 1991. *Randomization and Monte Carlo methods in biology*. Chapman and Hall, London.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220.
- Oden, N.L. & Sokal, R.R. 1992. An investigation of three-matrix permutation tests. *J. Class.* 9: 275-290.
- Orlóci, L. 1967. An agglomerative method for classification of plant communities. *J. Ecol.* 55: 193-205.
- Orlóci, L. 1978. *Multivariate analysis in vegetation research*. 2nd ed. Junk, The Hague.
- Orlóci, L. 1993. The complexities and scenarios of ecosystem analysis. In: Patil, G.P. & Rao, C.R. (eds.) *Multivariate environmental statistics*, pp. 421-430. North Holland/Elsevier, New York, NY.
- Orlóci, L. & Beshir, E. 1976. A heuristic test for homogeneity in species populations. *Vegetatio* 31: 141-145.
- Orlóci, L. & Kenkel, N.C. 1985. *Introduction to data analysis; with examples from population and community ecology*. International Co-operative Publishing House, Fairland, MD.
- Orlóci, L. & Pillar, V.D.P. 1989. On sample size optimality in ecosystem survey. *Biométrie-Praximétrie* 29: 173-184.
- Orlóci, L., Feoli, E., Lausi, D. & Nimis, P. 1986. Estimation of character structure convergence (divergence) in plant communities; a nested hierarchical model. *Coenoses* 1: 11-2.
- Pillar, V.D.P. 1988. *Fatores de ambiente relacionados à variação da vegetação de um campo natural*. M.Sc. Dissertation, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.
- Pillar, V.D.P. & Orlóci, L. 1993a. *Character-based community analysis; The theory and an application program*. SPB Academic Publishing, The Hague.
- Pillar, V.D.P. & Orlóci, L. 1993b. Taxonomy and perception in vegetation analysis. *Coenoses* 8: 53-66.
- Pitman, E.J.G. 1937. Significance tests which may be applied to samples from any populations. II. The correlation coefficient. *J. R. Stat. Soc. Ser. B* 4: 225-232.
- Podani, J. 1991. On the standardization of Procrustes statistics for the comparison of ordinations. *Abstr. Bot.* 15: 43-46.
- Potvin, C. & Roff, D.A. 1993. Distribution-free and robust statistical methods: viable alternatives to parametric statistics. *Ecology* 74: 1617-1628.
- Quinn, J.F. & Dunham, A.E. 1983. On hypothesis testing in ecology and evolution. *Am. Nat.* 122: 602-617.
- Rosengurt, S.B., Arrillaga, B. & Izaguirre, P. 1970. *Gramineas Uruguayas*. Universidad de la Republica, Montevideo.
- Scheffé, H. 1953. A method for judging all contrasts in the analysis of variance. *Biometrika* 40: 87-104.
- Smouse, P.E., Long, J.C. & Sokal, R.R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35: 627-632.
- Sokal, R.R. & Rohlf, F.J. 1962. The comparison of dendrograms by objective methods. *Taxon* 11: 33-39.
- Sokal, R.R., Oden, N.L., Thomson, B.A. & Kim, J. 1993. Testing for regional differences in means: distinguishing inherent from spurious spatial autocorrelation by restricted randomization. *Geogr. Anal.* 25: 199-210.
- ter Braak, C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.
- ter Braak, C.J.F. 1994. Canonical community ordination. Part I: Basic theory and linear methods. *Ecoscience* 1: 127-140.
- ter Braak, C.J.F. & Wiertz, J. 1994. On the statistical analysis of vegetation change: a wetland affected by water extraction and soil acidification. *J. Veg. Sci.* 5: 361-372.
- van der Maarel, E. 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity. *Vegetatio* 39: 97-114.
- Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58: 236-244.

Received 6 February 1995;

Revision received 31 October 1995;

Accepted 29 January 1996;