Research

Oikos 119: 1674–1684, 2010
doi: 10.1111/j.1600-0706.2010.18334.x
© 2010 The Authors. Journal compilation © 2010 Oikos
Subject Editor: José Alexandre Felizola. Accepted 23 February 2010

Improving indicator species analysis by combining groups of sites

Miquel De Cáceres, Pierre Legendre and Marco Moretti

M. De Cáceres (miquelcaceres@gmail.com), P. Legendre and M. Moretti, Dépt de Sciences Biologiques, Univ. de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, H3C 3J7, Canada. Present address for MDC: Centre Tecnològic Forestal de Catalunya, Ctra. St. Llorenç de Morunys km 2, ES–25280 Solsona, Catalonia, Spain. MM also at: Swiss Federal Research Inst., WSL, Ecosystem Boundaries, Via Belsoggiorno 22 CH–6500 Bellinzona, Switzerland.

Indicator species are species that are used as ecological indicators of community or habitat types, environmental conditions, or environmental changes. In order to determine indicator species, the characteristic to be predicted is represented in the form of a classification of the sites, which is compared to the patterns of distribution of the species found at the sites. Indicator species analysis should take into account the fact that species have different niche breadths: if a species is related to the conditions prevailing in two or more groups of sites, an indicator species analysis undertaken on individual groups of sites may fail to reveal this association. In this paper, we suggest improving indicator species analysis by considering all possible combinations of groups of sites and selecting the combination for which the species can be best used as indicator. When using a correlation index, such as the point-biserial correlation, the method yields the combination where the difference between the observed and expected abundance/frequency of the species is the largest. When an indicator value index (IndVal) is used, the method provides the set of site-groups that best matches the observed distribution pattern of the species. We illustrate the advantages of the method in three different examples. Consideration of combinations of groups of sites provides an extra flexibility to qualitatively model the habitat preferences of the species of interest. The method also allows users to cross multiple classifications of the same sites, increasing the amount of information resulting from the analysis. When applied to community types, it allows one to distinguish those species that characterize individual types from those that characterize the relationships between them. This distinction is useful to determine the number of types that maximizes the number of indicator species.

Indicator species are species that, due to their niche preferences, can be used as ecological indicators of community types, habitat conditions, or environmental changes (McGeoch 1998, Carignan and Villard 2002, Niemi and McDonald 2004). They are usually determined using an analysis of the relationship between the observed species presence–absence or abundance values in a set of sampled sites and a classification of the same sites (Dufrêne and Legendre 1997). Depending on the objective of the study, the groups of sites in the classification may represent different qualitative characteristics of the ecosystem, such as habitat or community types, environmental or succession states, or the levels of controlled experimental designs. Since indicator species analysis relates two elements, the species and the groups of sites, it can be used for gaining information on either or both. Indeed, indicator species analysis allows the characterization of the qualitative environmental preferences of the target species (for instance, when the groups are habitat types), and identifies indicators of particular groups of sites, which can be used in further surveys. The applications of indicator species analysis are many, including conservation, land management, landscape mapping, or design of natural reserves. Indicator species are commonly referred to as 'diagnostic species' in vegetation studies (Chytrý et al. 2002).

Indicator species analysis should take into account the fact that niche breadths vary among species. This means that some species may be related to one group of sites, while others may be related to more than one group (Tsiripidis et al. 2009). European phytosociologists were aware of this when they determined indicator species of groups belonging to different hierarchical levels of a vegetation classification (Barkman 1989). For example, some species could be indicators of all types of calcareous grasslands whereas others would point to dry calcareous grasslands on slopes with shallow soils. The fact that phytosociologists adopted a hierarchical classification scheme for communities influenced the development of numerical tools like the well-known two-way indicator species analysis (TWINSPAN, Hill 1979). This procedure takes a site-by-species data table and performs a hierarchical classification of the sites and, at the same time, determines the indicator species for the two sides of each split in the hierarchy. The main problem of TWINSPAN for indicator species analysis is that its multivariate nature makes the indicator value of one species dependent on the abundances of the remaining species in the data table (McGeoch and Chown 1998). Dufrêne and Legendre (1997) introduced the indicator value (IndVal) method, which treats each species separately. Moreover, unlike TWINSPAN, which generates the classification of the sites used in the later steps of the analysis, the IndVal method uses an already existing partition of the sites as input. The way this partition is defined is left to the user. It may have been obtained on the basis of environmental variables, community composition, or otherwise. The IndVal method determines the group of sites, among those forming the partition, to which the target species is most strongly related, as measured by the indicator value index. De Cáceres and Legendre (2009) recently developed a framework of statistical indices that can be used to measure the degree of association between the target species and each group of sites: we will hereafter use the word 'association' to refer to this kind of relationships. In order to cope with the different niche breadths of the species, Dufrêne and Legendre (1997) suggested to compare the indicator values derived from different partitions of the data, each corresponding to a different hierarchical level.

A hierarchical classification of sites may be too rigid to capture the niche preference of the target species. This is so because hierarchies do not consider all possible combinations of low-level clusters of the hierarchy, but only those allowed by the nested topology. For example, an indicator species analysis based on a hierarchical classification will be inadequate for a target species preferring two low-level clusters that do not form a node in the immediate upper clustering level. Acknowledging this fact, we present here an improvement of the IndVal approach, which consists in considering all structures at higher levels that come from combinations of groups of the initial partition of sites. For each species, the combination of site groups to be retained and tested for statistical significance is the one with the maximum association strength. The consideration of combinations of groups of sites presents some advantages. With respect to the species, it provides a fine characterization of the species habitat preferences from qualitative environmental data. With respect to the groups of sites, it allows one to distinguish between those species that characterize individual groups and those that explain the similarity between individual groups because they are associated to some larger grouping of the sites.

In the following section, we explain the method of indicator species analysis with combinations of groups in a detailed way. We consider two types of association indices (De Cáceres and Legendre 2009), the indicator value and the correlation value, and explain their differences in relation to the combinatorial approach. We then illustrate the advantages of the method using three different examples. First, we re-analyze the tree species-habitat associations found in Barro Colorado Island (Harms et al. 2001). Second, we determine beetle indicator species in response to fire, by crossing two classifications of the sites in the indicator species analysis. Third, we show how our approach can be helpful when using the number of indicator species to determine the optimal number of community types present in a community data table.

Indicator species analysis with combinations of site groups

The method

The method of indicator species analysis with combinations of site groups is an extension of the original IndVal method (Dufrêne and Legendre 1997). The input consists in two elements: the target species data vector, containing occurrence or abundance values at locations or sites, and a partition of the sites into a set of k non-overlapping classes, hereafter called 'site groups' (step 1 in Fig. 1). The main difference between the original IndVal method and the extension presented here is the following: whereas the original method considers the association between the target species and each of the k site groups in the partition, the extended method considers the association between the target species and each of the possible groups of sites that arise from combining (union operation) the site groups (step 2 in Fig. 1). If there are k different site groups in the partition, the number of possible combinations is 2^k–1, including the set of all sites. The original IndVal method looked for the group of sites to which the species was maximally associated. Analogously, the site-group combination to be retained in our method is the one showing the strongest association with the target species (step 3). Of course, the maximum association value may differ depending on the association index used (De Cáceres and Legendre 2009). In the next subsection, we present two types of association indices (Table 1) and discuss how the method of indicator species analysis with combinations of site groups differs when one or the other type is used. Before reporting the target species as indicator of the selected site-group combination, the association must be tested for statistical significance (step 4). We explain in subsection 'Statistical inference' how the significance test is carried out in the original IndVal and in the extended method.

Association indices

Indices for assessing the strength of association between species and groups of sites are reviewed in De Cáceres and Legendre (2009). We discuss here two indices that we modified to deal with combinations of site groups:

1) 'Indicator value' indices are specially designed to assess the predictive value of a species as indicator of a combination of site groups. The indicator value ($\text{IndVal}_{\text{ind}}$ in Table 1) index (from Dufrêne and Legendre 1997) is calculated as the product of two quantities, called A and B. Quantity A gives the probability of a site being a member of the site-group combination when the species has been found at that site (i.e. the positive predictive power of the species as ecological indicator of the site-group combination, Murtaugh 1996). Quantity B informs of how frequently (and hence how easily) the species is found at sites of the site-group combination under study.
2) 'Correlation indices' assess the positive or negative preference of the species for the environmental conditions prevailing within sites belonging to the site-group combination, compared to the remaining sites. The phi coefficient of association is a correlation index commonly used in vegetation studies to identify diagnostic species for plant community types (Chytrý et al. 2002). In the case of species abundance data, De Cáceres and Legendre (2009) suggest the point-biserial correlation coefficient (r_{pb} in Table 1). In our case, r_{pb} is equal to the Pearson correlation between a (binary) variable indicating whether the site belongs to the site-group combination under study, or not, and a (quantitative) variable containing the abundance of the species.

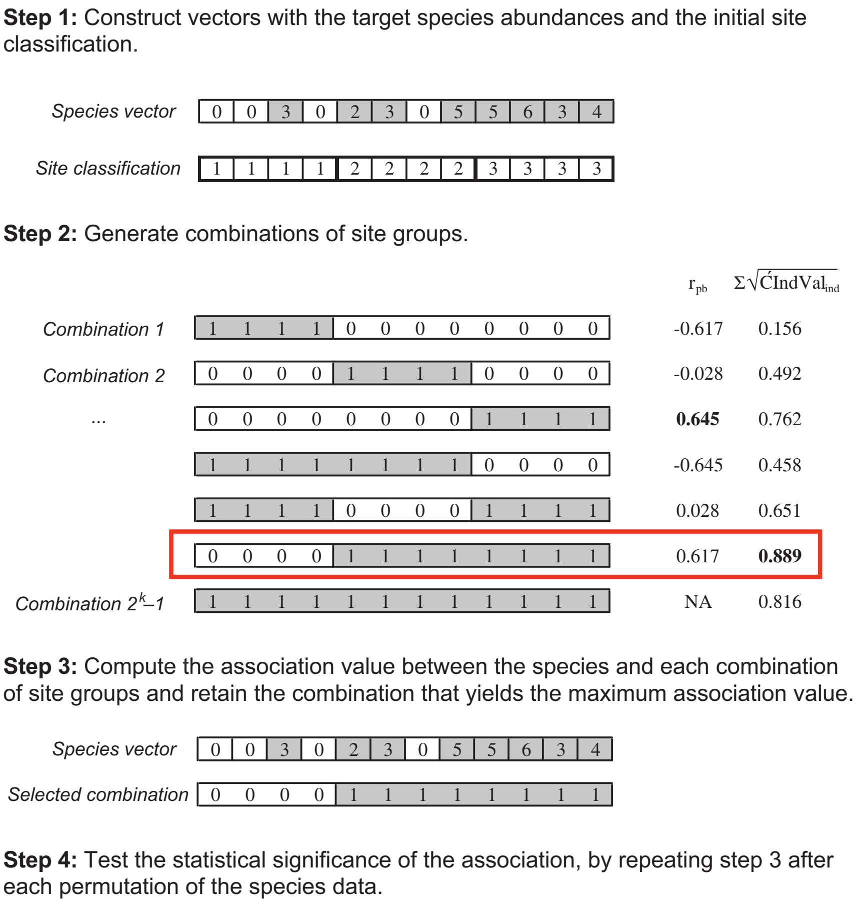

Step 1: Construct vectors with the target species abundances and the initial site classification.
Species vector 0 0 3 0 2 3 0 5 5 6 3 4
Site classification 1 1 1 1 2 2 2 2 3 3 3 3
Step 2: Generate combinations of site groups.
r_pb ΣĆIndVal_ind
Combination 1 1 1 1 1 0 0 0 0 0 0 0 0 -0.617 0.156
Combination 2 0 0 0 0 1 1 1 1 0 0 0 0 -0.028 0.492
... 0 0 0 0 0 0 0 0 1 1 1 1 0.645 0.762
1 1 1 1 1 1 1 1 0 0 0 0 -0.645 0.458
1 1 1 1 0 0 0 0 1 1 1 1 0.028 0.651
0 0 0 0 1 1 1 1 1 1 1 1 0.617 0.889
Combination 2^k–1 1 1 1 1 1 1 1 1 1 1 1 1 NA 0.816
Step 3: Compute the association value between the species and each combination of site groups and retain the combination that yields the maximum association value.
Species vector 0 0 3 0 2 3 0 5 5 6 3 4
Selected combination 0 0 0 0 1 1 1 1 1 1 1 1
Step 4: Test the statistical significance of the association, by repeating step 3 after each permutation of the species data.

Figure 1. Description of the steps involved in indicator species analysis with site-group combinations. Sites occupied by at least one individual of the target species and site memberships to the site-group combinations are indicated in gray.

Indicator species analysis may give different results depending on the association index used (see example in Fig. 1). Indicator value indices do not consider species absences outside the site-group combination under study (De Cáceres and Legendre 2009). As a result, when computed across site-group combinations, the combination that gets the highest association value is the one that best matches the observed presences of the species. Indicator species analysis using indicator value indices is a good tool to simplify and understand an observed species pattern. A randomly distributed species will have higher indicator values for larger site-group combinations (see expected values for random species in Table 1). As a consequence, randomly distributed species will tend to be matched with the set of all sites; that association cannot be tested for significance. Unlike indicator values, correlation indices take into account species absences both inside and outside the site-group combination under study. The set of all sites cannot be considered in a study involving correlation indices; the number of possible combinations is therefore $2^k–2$. An indicator species analysis using a correlation index will look for the site-group combination that has the highest difference between the species observed and expected frequency (or abundance). We provide in the Supplementary material Appendix 1 the results of a small simulation study; it may prove useful to those readers interested in better understanding the differences between r_{pb} and $IndVal_{ind}$.

Readers of Table 1 will notice that we present two versions of $IndVal_{ind}$ and r_{pb}. The standard (non-equalized) indices give the same weight to all individual sites; they are appropriate when the number of sites each site group is proportional to the ecological variability of the group (i.e. the larger the number of sites, the larger the ecological variability of the site group), but they may produce inaccurate estimates if some site groups with similar variability are over-sampled with respect to others (Tichý and Chytrý 2006). In order to avoid the potential problem of unbalanced sampling, group-equalized indices have to be used (Tichý and Chytrý 2006, De Cáceres and Legendre 2009). Group-equalized indices give equal weights to all site groups, therefore assuming that all have the same ecological variability. Under the present context of site-group combinations, group-equalized indices assume that the ecological variability of each site-group combination is proportional to the number of site groups it contains. Group-equalized indices will be our choice throughout this paper.

Statistical inference

In order to report that a target species is associated to a site-group or a site-group combination, we first need to reject the null hypothesis that negates this association. A permutation test is a procedure that involves comparing an observed test statistic with a distribution obtained by randomly reordering (i.e. permuting) the data. De Cáceres and Legendre (2009) describe a permutation test to assess the statistical

Table 1. Mathematical formulae and expected values of the indicator value ($\text{IndVal}_{\text{ind}}$) and the point biserial correlation (r_{pb}) indices when applied to combinations of site groups. Both non-equalized and group-equalized versions are presented.

		Indicator value	Point biserial correlation
Non-equalized	index	$\sqrt{\text{IndVal}_{\text{ind}}} = \sqrt{A_{\text{ind}} \times B} = \sqrt{\frac{a_C}{a} \times \frac{n_C}{N_C}}$	$r_{pb} = \frac{N \times a_C - a \times N_C}{\sqrt{(N \times l^2 - a^2) \times (N \times N_C - N_C^2)}}$
	expected value	$\sqrt{\frac{N_C}{N} \times \frac{n}{N}}$	0
Group-equalized	index	$\sqrt{\text{IndVal}^g_{\text{ind}}} = \sqrt{A^g_{\text{ind}} \times B} = \sqrt{\frac{a^g_C}{a^g} \times \frac{n_C}{N_C}}$	$r^g_{pb} = \frac{N \times a^g_C - a^g \times N^g_C}{\sqrt{(N \times l^{g2} - a^{g2}) \times (N \times N^g_C - N^{g2}_C)}}$
	expected value	$\sqrt{\frac{c}{k} \times \frac{n}{N}}$	0

Notation (following De Cáceres and Legendre 2009 and previous works): Let N be the number of sites in the data set, N_i be the number of sites belonging to site group i, n the number of sites where the target species occurs, and ni the number of sites in site group i where it occurs. Let then a be the sum of abundances of the target species over all sites, a_i the sum of its abundances in site group i, l the norm of the vector abundances of the target species, and l_i the norm of the abundance vector of the target species in site group i. Moreover, let K be the set of all k site groups and C be a set of c site groups conforming a particular site-group combination. We define$n_C = \sum_{i \in C} n_i$ and $a_C = \sum_{i \in C} a_i$ as the sum of occurrences and abundances of the species in the site-group combination C, respectively; and $N_C = \sum_{i \in C} N_i$ for the number of sites belonging to it. For group-equalized indices, the following modified quantities have to be defined:$N^g_C = \frac{N_C}{k}$, $a^g_C = \frac{N}{k}\sum_{i \in C}(a_i / N_i)$, $a^g = \frac{N}{k}\sum_{i \in K}(a_i / N_i)$ and$l^{g2} = \frac{N}{k}\sum_{i \in K}(l_i^2 / N_i)$.

significance of the association between the target species and a given group of sites. If the null hypothesis of no association is true, the association value computed after randomly reassigning species occurrence or abundance values to sites will be similar or very close to that observed for the original, unpermuted data. The p-value of the permutation test for positive (negative) species preference is the proportion of the permutations that yielded the same or higher (lower) association values than that observed for the unpermuted data.

When considering site-group combinations, permutation tests can be performed using the indices in Table 1 as test statistics. The permutation test described above is valid for assessing the statistical significance of the association between the target species and *a given* site-group combination. However, when assessing the significance of the association between the target species and the site-group combination with maximum association value, there is a caveat in the permutation test. The process of selecting the combination with maximum association value (step 3 in Fig. 1) increases the probability of finding a significant result, because the selected site-group combination is not independent of the species pattern. This means that random species will be found significant more often than the significance level, and the permutation test described above will not have a correct level of type I error. In order to solve this problem, we need to redo the selection process after each permutation of the species data. This modified permutation test (step 4 in Fig. 1) uses the maximum association value as the test statistic. It will have a correct level of type I error for random species, because we incorporate the process of selecting the site-group combination into the distribution of the null hypothesis of no association. Note that Dufrêne and Legendre (1997) also used the maximum IndVal, among the individual site groups, as the statistic for their permutation test. Therefore, the difference again consists in that we consider combinations of site groups instead of individual groups only.

When reporting the results of indicator species analysis for several species, users should be aware of multiple testing issues. Let us say that we conduct indicator species analysis with $\alpha = 0.05$ on 200 species without correcting for multiple testing and obtain 10 significant results. If we say that there are 10 indicator species, this 'experiment-wise' statement will probably be wrong, because under the null hypothesis of no association the expected number of significant results with 200 species is $\alpha \times 200 = 10$. Corrections for multiple testing are advisable in this case. These procedures (e.g. see the 'p.adjust' function of R) modify the p-values in order to keep the probability of finding, among all the statistical tests, at least one significant result at the chosen significance level. After the correction, we can report the number of significant indicators more safely. If, however, we are interested in reporting that a given species is an indicator, we do not need any correction because we are not making any experiment-wise statement. If the test is exact, the probability of type I error will be equal to α in that case.

Real data examples

Example 1. Species–habitat associations in Barro Colorado Island

One of the possible applications of indicator species analysis is to determine the ecological preference of species among a given set of habitat types. We re-examine here the habitat-association patterns of trees and shrubs within the 50-ha permanent Forest Dynamics plot of Barro Colorado Island (BCI), Panama. Habitat-association patterns in BCI were first studied by Harms et al. (2001) from a seven-habitat classification of 20 × 20 m grid cells in the plot based on the combination of topography, hydrology and historical facts. The seven habitat types are: high plateau (H, flat areas above 152 m); low plateau (L, flat areas below 152 m); areas with more than 7° in slope (S), which, due to the geology and hydrology of BCI, are moister than the plateaus later into the dry season; streams flanked with steep ravines (R); a seasonally-inundated swamp (W); secondary (about 100 years old) forest (F); and mixed habitats of difficult classification (M). Even if some species have an optimum among the seven original habitats, many of them may be more strongly related to two or more habitats types. We therefore intend to provide a more accurate characterization of species habitat niche preferences by considering combinations of habitat types.

We took the first BCI tree census, carried out in 1982–83, and counted the number of individuals of each tree species (i.e. the number of live stems with at least 1 cm dbh) in 20 × 20 m grid cells. We assessed the strength of the associations with the (group-equalized) r_{pb} index because the original study of Harms et al. (2001) aimed at determining both the positive and negative species habitat preferences. We compared the number of species with significant habitat preferences in the combinatorial analysis (there were $2^7–1 = 127$ combinations) to the same number obtained in the analysis performed considering the seven individual habitats only. In both cases, we determined the statistical significance of maximum association value using the permutation test described above. In order to prevent inflation of type I error due to autocorrelation in the unpermuted data, we restricted the permutations to those allowed by a toroidal shift (Lotwick and Silverman 1982). Briefly, permutations were obtained by using a two-dimensional torus connecting the map margins and then sliding one variable map (i.e., the target species abundance values) over the other (i.e. the memberships to the site-group combination).

Among the 307 tree species, 64 (21%) showed a significant association in the indicator value analysis considering habitat combinations, whereas if habitat combinations were not considered 44 (14%) species showed significant results. The low proportion of species showing preferences may be due to the small amount of topographic variation in the BCI plot. Among the 64 species showing a significant habitat preference, 34 were associated to a single habitat type, 17 were associated to the combination of two habitats, 10 were associated to three or four habitats, and three species were associated to five or six habitats (Table 2). Harms et al. (2001) mentioned a group of species strongly associated with slope areas, containing *Beilschmiedia pendula*, *Chrysochlamys eclipes, Poulsenia armata, Unonopsis pittieri* and *Virola surinamensis*. According to the results of the indicator species analysis with habitat combinations, only *U. pittieri* is associated to slopes only (Table 2). The remaining four species have a broader habitat preference. The habitats preferred by *B. pendula* are slopes and the low plateau. *C. eclipes* and *P. armata* are strongly associated to slopes and streamside habitats. *V. surinamensis* is weakly associated with slopes, streamside and swamp habitats. Knowing that these three habitats retain more moisture during the dry season makes this result consistent with Fisher et al. (1991), who wrote that *V. surinamensis* can persist and grow in shaded understory as long as there is no strong dry season. When habitat combinations were not considered, the habitat with maximum association for these four species was slopes but the statistical significance was lower (Table 2). Therefore, whereas Harms et al. (2001) correctly stated that those five species grow preferably on slopes, we achieved here a more detailed determination of the niche preferences of these species along the moisture gradient. From the results presented in Table 2, one can conclude that if the target species is actually associated to more than one habitat, the high frequency or abundance in one of them can hinder finding statistically significant results for another habitat. The combinatorial approach also allowed detecting some species that were positively associated to several habitats, meaning that they were avoiding specific habitats. For example, *Hybanthus prunifolius* (Violaceae) is found everywhere in BCI except in the moister habitats, i.e. the swamp (W) and near the streams (R). One may say that this species is negatively associated to these moister habitats.

Example 2. Indicator beetle species responding to fire and altitude

This application will show how indicator species analysis with combination of site groups can be applied to interpret the results of field mensurative pseudo-experiments with crossed factors. This is achieved by using as input the classification of sites derived from crossing the factors. Beetles are frequently used to indicate habitat alteration (Rainio and Niemelä 2003, Moretti et al. 2004). In this application, we identify the post-fire beetle species responses along an altitudinal gradient. The data are part of a broader research project which started in 2004, one year after a wildfire destroyed 300 ha of conifer forest between 800 and 2200 m above sea level (m a.s.l.) in the Swiss Alps (Wohlgemuth et al. 2006). We used the site classification derived from a two-way sampling design that consisted of three fire treatments (i.e. unburnt area, margin of the burnt area, center of the burnt area) across three altitudes (i.e. 1200, 1450, 1700 m a.s.l.). Longhorn beetles (Cerambycidae) and metallic wood-boring beetles (Buprestidae) were sampled weekly between April and September 2005 and 2006 (two and three years after the fire) at 18 trap sites along three transects at different altitudes, by means of one pitfall trap and one window trap per trap site. Specimens were identified at the species level, and the number of individuals per species counted. Using a canonical redundancy analysis biplot, we found that most of the species were associated to the margin and center burnt sites (Moretti et al. 2010). Our aim in this example of application was to use indicator species analysis to give a more detailed description of the post-fire response of each beetle species. Since we wanted to describe the pattern of each species, we used the group-equalized $\text{IndVal}_{\text{ind}}$ as association index. We carried out three indicator species analyses on log-transformed individual counts. In the first analysis, we used the three fire treatment levels as the classification of sites (which yields $2^3–1 = 7$ combinations of levels). The second analysis was conducted using the three altitude levels (seven combinations again). We ran the third analysis using the nine conditions

Table 2. Results of indicator species analysis on the BCI vegetation data. For each of the 64 species, we indicate the habitat combination that obtained the highest correlation (Hab. Comb.), the value of the correlation (r_{pb}), and the statistical significance of the association (p-value). For patterns containing more than one habitat, we also indicate the results obtained for the most highly related single habitat when combinations are not considered (Hab.). Habitat codes are defined in the main text.

Species	Hab. Comb.	r_{pb}	p	Species	Hab. Comb.	r_{pb}	p	Hab.	r_{pb}	p
34 species associated to one habitat				17 species associated to two habitats						
Alseis blackiana	F	0.635	0.002	*Trichilia tuberculata*	F+H	0.452	0.030	F	0.360	0.141
Gustavia superba	F	0.635	0.003	*Ficus yoponensis*	F+R	0.189	0.009	R	0.157	0.038
Guettarda foliacea	F	0.399	0.004	*Hieronyma alchorneoides*	F+W	0.262	0.003	W	0.185	0.036
Adelia triloba	F	0.375	0.024	*Ficus costaricana*	H+W	0.159	0.047	W	0.151	0.078
Casearia aculeata	F	0.275	0.004	*Lacistema aggregatum*	L+R	0.319	0.010	R	0.260	0.046
Pseudobombax septenatum	F	0.275	0.020	*Beilschmiedia pendula*	L+S	0.379	0.012	S	0.274	0.064
Guazuma ulmifolia	F	0.188	0.027	*Tachigali versicolor*	L+S	0.365	0.003	S	0.245	0.062
Piper aequale	R	0.389	0.003	*Cassipourea elliptica*	M+W	0.290	0.049	W	0.246	0.125
Nectandra purpurea	R	0.258	0.007	*Poulsenia armata*	R+S	0.516	0.002	S	0.434	0.017
Marila laxiflora	R	0.256	0.007	*Chrysochlamys eclipes*	R+S	0.437	0.010	S	0.307	0.092
Laetia procera	R	0.243	0.017	*Trophis caucana*	R+S	0.417	0.046	R	0.303	0.203
Brosimum alicastrum	R	0.229	0.011	*Conostegia cinnamomea*	R+S	0.327	0.006	R	0.251	0.035
Myrcia gatunensis	R	0.212	0.048	*Symphonia globulifera*	R+S	0.296	0.033	R	0.254	0.075
Conostegia bracteata	R	0.208	0.027	*Pentagonia macrophylla*	R+S	0.245	0.043	R	0.163	0.228
Cyathea petiolata	R	0.197	0.016	*Maquira guianensis*	R+S	0.231	0.042	R	0.199	0.071
Ocotea whitei	S	0.362	0.035	*Genipa americana*	R+W	0.194	0.044	W	0.155	0.095
Unonopsis pittieri	S	0.341	0.035	*Ficus insipida*	R+W	0.140	0.050	W	0.095	0.331
Bactris major	W	0.554	0.002							
Elaeis oleifera	W	0.484	0.002	10 species associated to three or four habitats						
Triplaris cumingiana	W	0.366	0.002	*Protium panamense*	L+R+S	0.475	0.030	R	0.286	0.162
Astrocaryum standleyanum	W	0.362	0.002	*Protium costaricense*	L+R+S	0.324	0.012	R	0.209	0.202
Attalea butyracea	W	0.327	0.003	*Annona acuminata*	M+R+W	0.236	0.045	W	0.206	0.082
Terminalia amazonia	W	0.317	0.023	*Inga goldmanii*	M+S+W	0.187	0.040	W	0.176	0.041
Tetrathylacium johansenii	W	0.290	0.003	*Virola surinamensis*	R+S+W	0.249	0.033	S	0.158	0.155
Tabebuia rosea	W	0.285	0.003	*Quararibea asterolepis*	F+H+R+S	0.352	0.004	F	0.266	0,029
Stylogyne turbacensis	W	0.264	0.003	*Swartzia simplex* var. *ochnacea*	F+L+R+S	0.403	0.025	F	0.381	0.021
Thevetia ahouai	W	0.227	0.028	*Mouriri myrtilloides*	H+L+M+R	0.357	0.039	L	0.284	0.149
Lindackeria laurina	W	0.227	0.049	*Desmopsis panamensis*	H+L+M+S	0.406	0.032	H	0.369	0.056
Piper reticulatum	W	0.223	0.013	*Hirtella triandra*	L+M+R+S	0.438	0.012	S	0.362	0.038
Piper colonense	W	0.210	0.015							
Inga laurina	W	0.191	0.021	3 species associated to five or six habitats						
Chrysophyllum cainito	W	0.175	0.039	*Hybanthus prunifolius*	F+H+L+M+S	0.503	0.002	F	0.495	0.003
Vismia macrophylla	W	0.169	0.022	*Inga sapindoides*	H+L+M+R+S	0.175	0.034	S	0.068	0.569
Maclura tinctoria	W	0.164	0.035	*Erythroxylum macrophyllum*	F+H+L+M+R+S	0.172	0.029	S	0.076	0.466

resulting from the crossing of treatment and altitude effects as classification of sites (which yields $2^9–1 = 511$ combinations). In all cases, the two survey years were considered to represent independent observations (giving four replicate observations of the beetle community in each area).

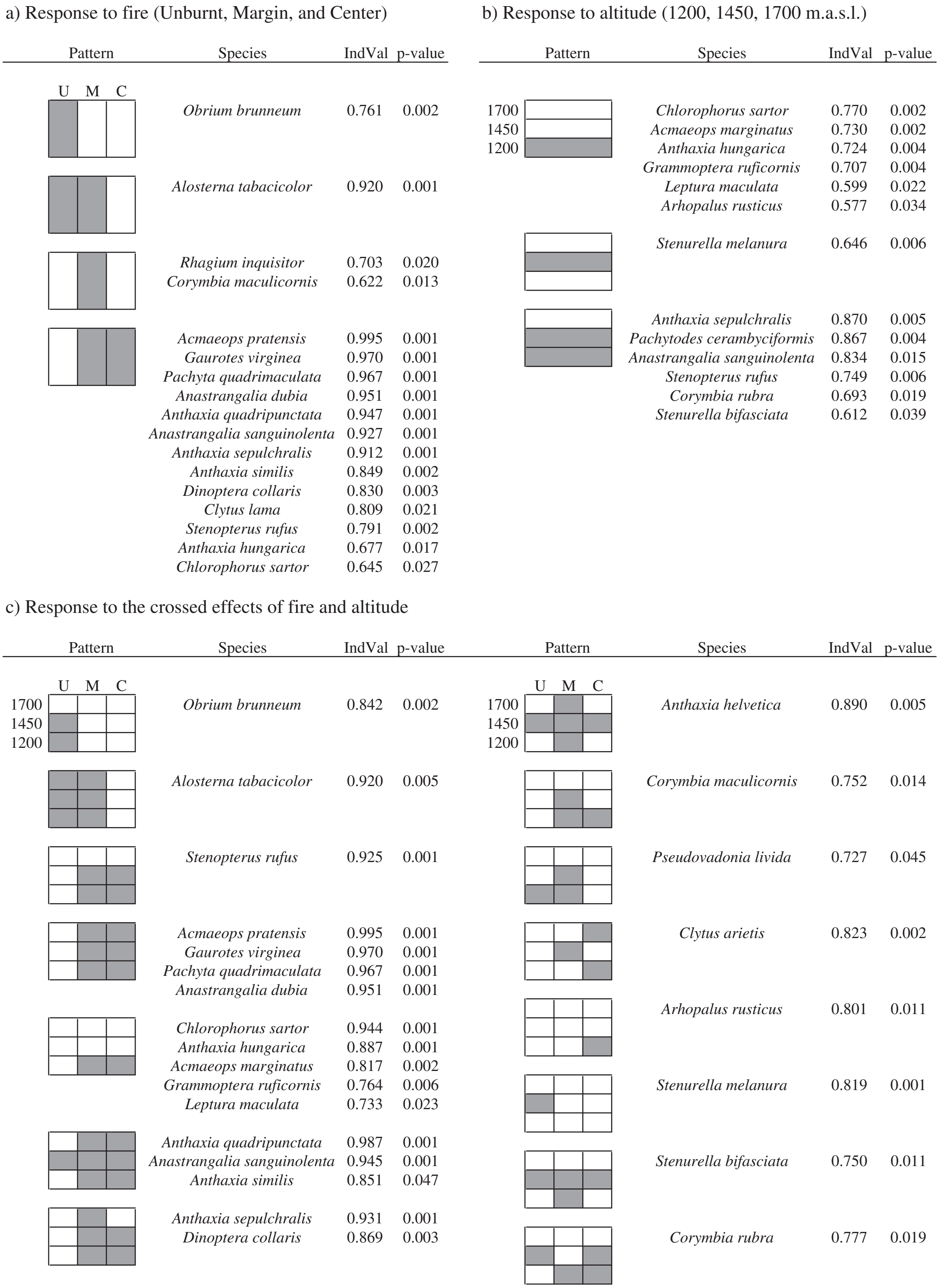
Among the 62 species available, indicator species analysis revealed 17 indicators responding to fire, most of them being associated to burnt areas (the combination of margin and center). The analysis on altitude belts yielded 13 indicators, mainly associated to low or low and intermediate altitudes. Five species were significant indicators in both analyses: *Chlorophorus sartor* and *Anthaxia hungarica* were at the same time indicators of burnt areas and low altitudes, whereas *Anthaxia sepulchralis, Anastrangalia sanguinolenta* and *Stenopterus rufus* were indicators of burnt areas and low and intermediate altitudes (Fig. 2a–b). The third analysis yielded 25 indicators; the higher number of indicators is likely to be due to the finer description of the environmental conditions of the sampled sites obtained by crossing the two factors (Fig. 2c). Most importantly, the third analysis clarified the situation that each species was indicating best. In this sense, we found the results of this third analysis to be a very good complement to the canonical biplot. For example, *Obrium brunneum* had been found to be an indicator of unburnt sites in the first analysis, but the third analysis revealed that its occurrence was restricted to low and intermediate altitudes. Using a minimum association value of $\sqrt{\text{IndVal}_{\text{ind}}} = 0.8$ as a threshold, four species were confirmed to be indicators of fire at all altitudes (*Acmaeops pratensis, Gaurotes virginea, Pachyta quadrimaculata, Anastrangalia dubia*). Three species, *Chlorophorus sartor, Anthaxia hungarica* and *Acmaeops marginatus*, were indicators of fire at low altitudes; while only *Stenopterus rufus* was a good indicator of fire in low and intermediate altitudes. Some species had higher number of individuals at the margin of the burnt area, like *Anthaxia helvetica*, while others, like *Alosterna tabacicolor*, occupied the burnt margin area coming from the forest, both benefiting from the positive margin effect. On the other hand, the two species appearing to be restricted to margin sites in the first analysis were not confirmed as good indicators in the third one. Thus, strict indicators of the margin sites are unlikely. Many of the indicator species found in our study are of great importance from the point of view of conservation, while several of the indicator species related to the burnt sites at different altitudes are rare in Switzerland and the information about their ecology is scarce. The combinatorial approach obtains detailed information about the ecological niche of species with relatively narrow habitat and microclimatic requirements.

Example 3. Determining the number of types in a community data table

Delimiting community types is a way to simplify the complexity of community data, providing a complementary view to non-canonical ordination methods (Legendre and Legendre 1998). Although they are artificial, community types may become very useful for the synthesis of ecological survey results. For instance, plant community types are often used to generate vegetation maps, but they can also be useful to characterize community-level successional changes over time. Arbitrary choices of a clustering method and the number of groups are usually a source of concern among the practitioners of clustering methods (Mucina 1997). The strategy normally employed consists in comparing different possible classifications using a pre-specified criterion. Many criteria exist, and most of them are based on the geometrical compactness and isolation of clusters (Milligan and Cooper 1985), but for community ecology data authors have suggested using the number of indicator species (Dufrêne and Legendre 1997, Tuomisto et al. 2003, Aho et al. 2008). If we take into account that community types are not real entities, and that community data is essentially of continuous nature, finding 'natural' or 'data-driven' community types is a certainly difficult, if not impossible, task. We do not pretend here to perform a comparison of criteria to determine the optimal classification of community data (Milligan and Cooper 1985, Aho et al. 2008), but rather to show the added value of considering site-group combinations when using indicator species as a criterion.

We took three community data sets: 1) the oribatid mite data of the *Sphagnum* mosses of lac Geai (Saint-Hippolyte, Québec, Canada) collected by Daniel Borcard and used in a number of publications (Borcard et al. 1992). The data set consists in individual counts of 35 mite species in 70 soil cores from a sampling area of 2.5 m by 10 m in size; it is available in the 'vegan' and 'ade4' R language libraries, for example; 2) wetland vegetation data from the alluvial plain of Adelaide River (Australia) studied by Bowman and Wilson (1986). The community data table is composed of 40 sites and 33 vascular plants whose abundance values are recorded following cover classes. It is available in the R language library 'indicspecies' (De Cáceres and Legendre 2009); 3) the BCI plot data, presented in our first application, consists in stem counts of 307 tree species in 1250 grid cells. We ran K-means partitioning (MacQueen 1967) on each of the three data sets after transforming the community data through the Hellinger transformation (Legendre and Gallagher 2001). This transformation decreases the importance of abundance over occurrence (as in the square root transformation) and avoids the double-zero problem when comparing species composition between sites (Legendre and Legendre 1998). In order to determine the best partition of each community data table, we first conducted several runs of K-means using different numbers of clusters, from two to eight. We then carried out four indicator species analyses on each data table and each partition. The four analyses come from the crossing of whether site-group combinations are considered or not, and whether $\text{IndVal}_{\text{ind}}$ or r_{pb} is used as the association index. We did not use the Hellinger transformation for the indicator species analyses in order to keep the results for each species independent. The mite data were log-transformed prior to indicator value analysis in order to reduce the influence of very large values. In the results of each analysis, we counted the total number of significant indicator species (after 999 permutations, $\alpha = 0.05$). In the analyses considering site-group combinations, we also counted how many, among the overall significant species, were associated to a single group.

For the oribatid mite data (Fig. 3a), the number of $\text{IndVal}_{\text{ind}}$ indicator species with and without combinations had a maximum (25–26 species) at $k = 4$, while the plot for r_{pb} was almost flat, indicating a lack of clear clustering structure. The number of species associated to single groups in the $\text{IndVal}_{\text{ind}}$ group-combination analysis was highest at $k = 2$



a) Response to fire (Unburnt, Margin, and Center)

Pattern	Species	IndVal	p-value
U M C ■ □ □	*Obrium brunneum*	0.761	0.002
■ ■ □	*Alosterna tabacicolor*	0.920	0.001
□ ■ □	*Rhagium inquisitor*	0.703	0.020
	Corymbia maculicornis	0.622	0.013
□ ■ ■	*Acmaeops pratensis*	0.995	0.001
	Gaurotes virginea	0.970	0.001
	Pachyta quadrimaculata	0.967	0.001
	Anastrangalia dubia	0.951	0.001
	Anthaxia quadripunctata	0.947	0.001
	Anastrangalia sanguinolenta	0.927	0.001
	Anthaxia sepulchralis	0.912	0.001
	Anthaxia similis	0.849	0.002
	Dinoptera collaris	0.830	0.003
	Clytus lama	0.809	0.021
	Stenopterus rufus	0.791	0.002
	Anthaxia hungarica	0.677	0.017
	Chlorophorus sartor	0.645	0.027

b) Response to altitude (1200, 1450, 1700 m.a.s.l.)

Pattern	Species	IndVal	p-value
1700 □ 1450 □ 1200 ■	*Chlorophorus sartor*	0.770	0.002
	Acmaeops marginatus	0.730	0.002
	Anthaxia hungarica	0.724	0.004
	Grammoptera ruficornis	0.707	0.004
	Leptura maculata	0.599	0.022
	Arhopalus rusticus	0.577	0.034
□ ■ □	*Stenurella melanura*	0.646	0.006
□ ■ ■	*Anthaxia sepulchralis*	0.870	0.005
	Pachytodes cerambyciformis	0.867	0.004
	Anastrangalia sanguinolenta	0.834	0.015
	Stenopterus rufus	0.749	0.006
	Corymbia rubra	0.693	0.019
	Stenurella bifasciata	0.612	0.039

c) Response to the crossed effects of fire and altitude

Pattern	Species	IndVal	p-value
U M C 1700 □ □ □ 1450 ■ □ □ 1200 ■ □ □	*Obrium brunneum*	0.842	0.002
■ ■ □ ■ ■ □ ■ ■ □	*Alosterna tabacicolor*	0.920	0.005
□ □ □ □ ■ ■ □ ■ ■	*Stenopterus rufus*	0.925	0.001
□ ■ ■ □ ■ ■ □ ■ ■	*Acmaeops pratensis*	0.995	0.001
	Gaurotes virginea	0.970	0.001
	Pachyta quadrimaculata	0.967	0.001
	Anastrangalia dubia	0.951	0.001
□ □ □ □ □ □ □ ■ ■	*Chlorophorus sartor*	0.944	0.001
	Anthaxia hungarica	0.887	0.001
	Acmaeops marginatus	0.817	0.002
	Grammoptera ruficornis	0.764	0.006
	Leptura maculata	0.733	0.023
□ ■ ■ ■ ■ ■ □ ■ ■	*Anthaxia quadripunctata*	0.987	0.001
	Anastrangalia sanguinolenta	0.945	0.001
	Anthaxia similis	0.851	0.047
□ ■ □ □ ■ ■ □ ■ ■	*Anthaxia sepulchralis*	0.931	0.001
	Dinoptera collaris	0.869	0.003

Pattern	Species	IndVal	p-value
U M C 1700 □ ■ □ 1450 ■ ■ ■ 1200 □ ■ □	*Anthaxia helvetica*	0.890	0.005
□ □ □ □ ■ □ □ ■ ■	*Corymbia maculicornis*	0.752	0.014
□ □ □ □ ■ □ ■ ■ □	*Pseudovadonia livida*	0.727	0.045
□ □ ■ □ ■ □ □ □ ■	*Clytus arietis*	0.823	0.002
□ □ □ □ □ □ □ □ ■	*Arhopalus rusticus*	0.801	0.011
□ □ □ ■ □ □ □ □ □	*Stenurella melanura*	0.819	0.001
□ □ □ ■ ■ ■ □ ■ □	*Stenurella bifasciata*	0.750	0.011
□ □ □ ■ □ ■ □ ■ ■	*Corymbia rubra*	0.777	0.019

Figure 2. Indicator species for the beetle data ($\alpha = 0.05$, 999 perm.) in response to fire (a), altitude (b), or the crossed effect of fire and altitude (c). In grey: the combination of site groups that is most strongly related to the species pattern.

(20 species), but very low (four species) at k = 4, indicating that at this level most species are better associated to combinations of two or three groups rather than to single groups. A decision to divide the data set into two groups of sites would be in accordance with Legendre (2005), who found two significant species associations for this data. For the wetland vegetation data (Fig. 3b), the number of indicator species without considering combinations was not a

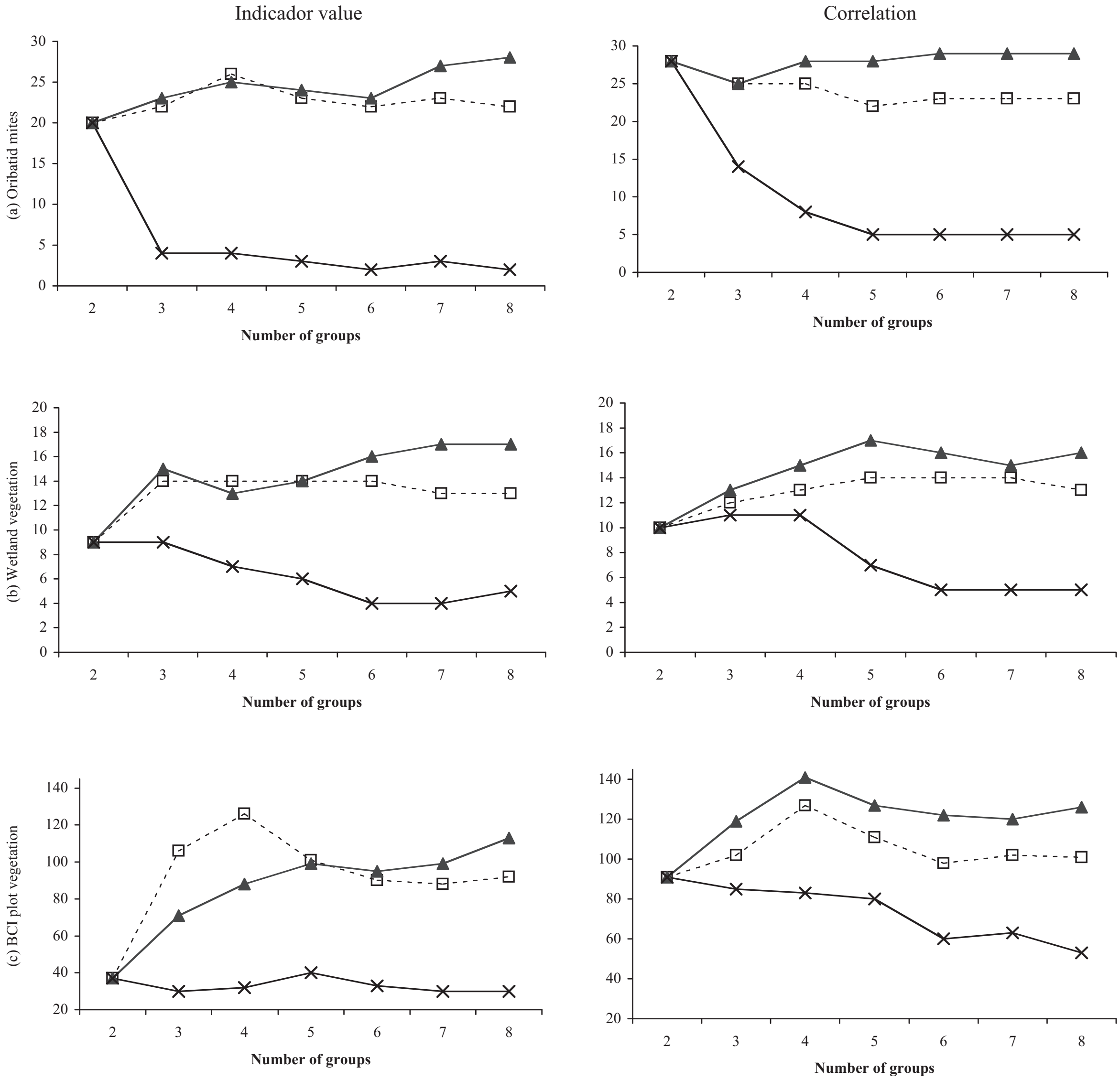


Figure 3. Number of statistically significant indicator species ($\alpha = 0.05$, 999 perm.) resulting from a classification using K-means partitioning with different numbers of groups on the three data sets (a–c) Empty squares: the number of indicator species when site-group combinations are not considered; triangles: the number of indicator species when site-group combinations are considered; crosses: the number of indicator species associated to a single site group when combinations are considered.

useful criterion, because the same number of indicators was obtained for many values of k. On the other hand, when site-group combinations were considered, there was a first maximum of 15 indicators at $k = 3$ for $\text{IndVal}_{\text{ind}}$ and at $k = 5$ for correlations. Finally, the number of indicators associated to individual groups pointed to $k = 2$ or 3 with $\text{IndVal}_{\text{ind}}$, and to $k = 3$ or 4 with r_{pb}. This last result is consistent with Dale (1988), who used the same data to compare fuzzy classification methods and pointed out that three or four groups of sites can be distinguished. The results in the case of BCI were surprising (Fig. 3c): there was a strong peak (126 species) at $k = 4$ without considering combinations. The fact that with combinations the number of $\text{IndVal}_{\text{ind}}$ indicator species was much lower reflects the fact that many of the BCI species occur in almost all 20 × 20 m cells of the plot. The combination analysis, using the indicator value index as the statistic, matched those species with the group combination that includes all sites. Since the statistical significance cannot be tested in that situation, the number of diagnostic species with significant combinations was lower. In the case of BCI, a partition into four or five site groups seems optimal and agrees with the habitats defined by Harms et al. (2001). Note that in the first application we found that only four among the seven habitat types had indicator species associated to them individually (Table 2).

Provided that one decides to compare classifications using the number of indicator species as a criterion, our results indicate that distinguishing between species associated to single groups and species associated to two or more groups can be helpful. While the indicators for single site groups effectively characterize the individual community types, the indicator species associated to two or more groups allow one to interpret the similarity between community types because they characterize a higher-level grouping structure. Nevertheless, a simulation study specifically focusing on this issue would be most useful.

Discussion

On considering site-group combinations in indicator species analysis

This paper stresses that indicator species analysis must take into account the fact that niche breadths vary among species. Tsiripidis et al. (2009) recently addressed the same issue, proposing an algorithm that considers the possible association of the species with more than one group of sites. Their method is based on the comparison between relative frequency values of the target species across site groups, using an arbitrary difference threshold. As a consequence, their approach does not allow statistical inference and only applies to presence-absence species data. In the present paper, we discussed a simple but powerful extension of the well-known IndVal method, based on exploring all possible combinations of site groups. In the application section we demonstrated several advantages of our approach compared to normal practice in indicator species analysis. For the benefit of interested readers, we included in the R language library 'indicspecies' (<http://sites.google.com/site/miquelde caceres/software>) (De Cáceres and Legendre 2009, supplement) a function that can be used to perform indicator species analysis with or without combinations.

Although the consideration of combinations of site-groups highly improves the flexibility of the indicator species analysis, some site-group combinations arising from the method may not be easy to interpret ecologically. Site-group combinations that are difficult to interpret can occur in at least three cases: 1) since the test uses a significance level (for example $\alpha = 0.05$), the reported association may in fact be a random event; 2) due to historical reasons (e.g. dispersal limitation), the reported association may reflect the pattern of the target species in the studied data set, but not its real ecological preference; 3) the sites forming the site-group combination may share some environmental characteristic(s) driving the species preference, that the ecologist was unaware of. Unless the investigator is interested in this last case, the value of the reported association will only be of practical usefulness when the indicated entity has received some interpretation. Fortunately, the method presented here allows restricting the site-group combinations to those that have ecological meaning according to the analyst. Previous applications of the original IndVal method were even more restrictive since they were assuming that the ecologically relevant combinations involved individual site-groups only.

On the limitations of indicator species analysis

Indicator species are useful albeit simple tools to indicate the qualitative state of an ecosystem when it is unknown. This means, for instance, that one can use indicator species to classify sites of unexplored ecosystems (e.g. in vegetation mapping) or to monitor the succession of a particular habitat or environmental gradient to natural and anthropogenic stressors over time (e.g. recover after pollution, disturbance by fire, climatic changes, etc). Depending on the purpose of the application, indicator species are categorized between environmental, ecological or biodiversity indicators (McGeoch 1998). Although the methods are very distinct from the ones used here, it is worth noting that diversity indicator species can be determined, for instance by selecting predictors of species richness in regression models (Mac Nally and Fleishman 2004). Besides its advantages, indicator species analysis also has several limitations. We will conclude by restating six of them, with the aim of preventing misuses and/or abuses of the method and, if possible, encourage future developments. (1) Although it provides a qualitative assessment of the target species niche, indicator species analysis is not the best method to describe species niches. There exist methods based on quantitative environmental data that better perform this task (Thuiller et al. 2004). (2) The spatial, temporal and environmental context of determination of indicator species is crucial and should always be mentioned and described (De Caceres et al. 2008, Willner et al. 2009). For example, the BCI species-habitat associations we studied here apply to the 50 ha BCI forest plot only. A broader topographical and geographical context of determination would have been necessary in order to state that the studied species indeed prefer the indicated habitats. Another example is the study of insect indicator species through time along a year: the results would be different from, and uncomparable with, the results of a study through space. (3) Users of indicator species analyses should bear in mind that they are trying to detect patterns of association, without knowing whether these patterns arise from the process attributed to them. For example, if the BCI species-habitat associations correspond to true niche preferences, one can expect that *U. pitteri* should grow and perform poorly outside the slope areas, and the four other species cited should be able to support relatively high growth and performance in a wider range of habitats. (4) More indicator species will be found than expected by chance when the classification of sites has been obtained from the species composition itself (De Cáceres and Legendre 2009). In this case, p-values must be taken with caution: they do not result from a genuine test of significance since the classification of sites is not independent from the species data used in the indicator species analysis. (5) IndVal's quantity B expresses how easily the target species is found at sites belonging to the site-group combination under consideration. The current calculation as a simple frequency does not allow distinguishing between the ability to detect the species in space and time (De Cáceres and Legendre 2009). (6) In the normal application of indicator species analysis, the presence of a species (or a group of species) is taken as an indication of the state of the ecosystem. IndVal's quantity A informs of the probability of the site-group combination given the fact that the species has been found. The accuracy of bioindication would be higher after performing probabilistic calculations from the detected presence of multiple species.

Acknowledgements – This research was funded by NSERC grant no. 7738 to P. Legendre. The BCI forest dynamics research project is part the Center for Tropical Forest Science, a global network of large-scale demographic tree plots. The beetle study is part of the Forest Dynamic program at the Swiss Federal Research Inst. WSL Birmensdorf leaded by Thomas Wohlgemuth. We thank Cinzia Pradella for the identification of beetle species, and Beat Wermelinger, Martin Obrist, Peter Duelli and Franco Fibbioli for their scientific and technical assistance. The authors feel especially grateful to Milan Chytry for constructive comments on previous versions of the manuscript.

References

Aho, K. et al. 2008. Using geometric and non-geometric internal evaluators to compare eight vegetation classification methods. – J. Veg. Sci. 19: 549–562.

Barkman, J. J. 1989. Fidelity and character-species, a critical evaluation. – Vegetatio 85: 105–116.

Borcard, D. et al. 1992. Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.

Bowman, D. M. J. S. and Wilson, B. A. 1986. Wetland vegetation pattern on the Adelaide River flood plain, Northern Territory, Australia. – Proc. R. Soc. Queensland 97: 69–77.

Carignan, V. and Villard, M. 2002. Selecting indicator species to monitor ecological integrity: a review. – Environ. Monitor. Assess. 78: 45–61.

Chytrý, M. et al. 2002. Determination of diagnostic species with statistical fidelity measures. – J. Veg. Sci. 13: 79–90.

Dale, M. B. 1988. Some fuzzy approaches to phytosociology. Ideals and instances. – Folia Geobot. Phytotax. 23: 239–274.

De Cáceres, M. and Legendre, P. 2009. Associations between species and groups of sites: indices and statistical inference. – Ecology 90: 3566–3574.

De Cáceres, M. et al. 2008. Assessing species diagnostic value in large data sets: a comparison between phi coefficient and Ochiai index. – J. Veg. Sci. 19: 779–788.

Dufrêne, M. and Legendre, P. 1997. Species assemblages and indicator species: the need for a flexible asymetrical approach. – Ecol. Monogr. 67: 345–366.

Fisher, B. L. et al. 1991. Survival and growth of *Virola surinamensis* yearlings: water augmentation in gap and understory. – Oecologia 86: 292–297.

Harms, K. E. et al. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. – J. Ecol. 89: 947–959.

Hill, M. O. 1979. TWINSPAN – a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. – Cornel Univ., NY.

Legendre, P. 2005. Species associations: the Kendall coefficient of concordance revisited. – J. Agric. Biol. Environ. Stat. 10: 226–245.

Legendre, P. and Gallagher, E. D. 2001. Ecologically meaningful transformations for ordination of species data. – Oecologia 129: 271–280.

Legendre, P. and Legendre, L. 1998. Numerical ecology (2nd english ed.). – Elsevier.

Lotwick, H. W. and Silverman, B. W. 1982. Methods for analysing spatial processes of several types of points. – J. R. Stat. Soc. Ser. B 44: 406–413.

Supplementary material (available online as Appendix O18334 at www.oikos.ekol.lu.se/appendix. Appendix 1. Numerical simulation study.

Mac Nally, R. and Fleishman, E. 2004. A successful predictive model of species richness based on indicator species. – Conserv. Biol. 18: 646–654.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observation. – In: LeCam, L. M. and Neyman, J. (eds), Proc. 5th Berkeley Symp. Math. Stat. Probabil. Univ. of California Press, pp. 281–297.

McGeogh, M. A. 1998. The selection, testing and application of terrestrial insects as bioindicators. – Biol. Rev. 73: 181–201.

McGeoch, M. A. and Chown. S. L. 1998. Scaling up the value of bioindicators. – Trends Ecol. Evol. 13: 46–47.

Milligan, G. W. and Cooper, M. C. 1985. An examination of procedures for determining the number of clusters in a data set. – Psychometrika 50: 159–179.

Moretti, M. et al. 2004. Arthropod biodiversity after forests fires: winners and losers in the winter fire regime of the southern Alps. – Ecography 27: 173–186.

Moretti, M. et al. 2010. Fire-induced taxonomic and functional changes in saproxylic beetle communities in fire sensitive regions. – Ecography doi.org/10.1111/j.1600-0587.2009.06172.x.

Mucina, L. 1997. Classification of vegetation: past, present and future. – J. Veg. Sci. 8: 751–760.

Murtaugh, P. A. 1996. The statistical evaluation of ecological indicators. – Ecol. Appl. 6: 132–139.

Niemi, G. J. and McDonald, M. E. 2004. Application of ecological indicators. – Annu. Rev. Ecol. Evol. Syst. 35: 89–111.

Rainio, J. and Niemelä, J. 2003. Ground beetles (Coleoptera: Carabidae) as bioindicators. – Biodiv. Conserv. 12: 487–506.

Thuiller, W. et al. 2004. Relating plant traits and species distributions along bioclimatic gradients for 88 *Leucadendron* taxa. – Ecology 85: 1688–1699.

Tichý, L. and Chytrý, M. 2006. Statistical determination of diagnostic species for site groups of unequal size. – J. Veg. Sci. 17: 809–818.

Tuomisto, H. et al. 2003. Floristic patterns along a 43-km long transect in an Amazonian rain forest. – J. Ecol. 91: 743–756.

Tsiripidis, I. et al. 2009. A new algorithm for the determination of differential taxa. – J. Veg. Sci. 20: 233–240.

Wohlgemuth, T. et al. 2006. Ecological resilience after fire in mountain forests of the central Alps. – In: Viegas, D. X. (ed.), V Int. Conf. Forest Fire Res., November 2006, Figueira da Foz, Portugal. ADAI/CEIF Univ. of Coimbra.

Willner, W. et al. 2009. Effects of different fidelity measures and contexts on the determination of diagnostic species. – J. Veg. Sci. 20: 130–137.