Niche Separation in Community Analysis: A New Method

Sylvain Doledec; Daniel Chessel; Clementine Gimaret-Carpentier

NICHE SEPARATION IN COMMUNITY ANALYSIS: A NEW METHOD

Sylvain Dolédec,[1] Daniel Chessel,[2] and Clémentine Gimaret-Carpentier[2]

[1]*UMR CNRS 5023, Laboratoire d'Ecologie des Hydrosystèmes Fluviaux, Université Lyon I, 69622*
Villeurbanne Cedex, France
[2]*UMR CNRS 5558, Laboratoire de Biomérie et Biologie Evolutive, Université Lyon I, 69622 Villeurbanne Cedex, France*

Abstract. The design and objective of a community study imply the selection of the appropriate ordination technique in terms of species response models and weighting options. In this paper, we start from the observation that existing two-table ordination techniques and related measures of niche breadth inevitably weight a sample in proportion to its abundance. We introduce a new multivariate method, which gives a more even weight to all sampling units, including those which are species poor or individual poor. We use this new method of analysis which we call OMI (for Outlying Mean Index) to address the question of niche separation and niche breadth. The Outlying Mean Index, or species marginality, measures the distance between the mean habitat conditions used by species (species centroid), and the mean habitat conditions of the sampling area (origin of the niche hyperspace), and OMI analysis places species along habitat conditions using a maximization of their mean OMI. Therefore, the position of the species depends on their niche deviation from a reference, which represents neither the mean nor the most abundant species, but a theoretical ubiquitous species that tolerates the most general habitat conditions (i.e., a hypothetical species uniformly distributed among habitat conditions). We demonstrate that OMI analysis is well suited for the investigation of multidimensional niche breadths in the case of strong limiting factors (e.g., meteorological conditions) or strong driving forces (e.g., longitudinal stream gradient). Furthermore, the analysis helps in finding which ecological factors are most important for community structure and organization and provides a separation of species based on their niche characteristics.

Key words: biotic interactions; multivariate analysis; niche breadth; niche position; outlying mean index (OMI); permutation test; sample weight; species–environment relationships.

INTRODUCTION

The niche concept as defined by Hutchinson (1957) is considered a cornerstone in ecology (Begon et al. 1996). This concept usually considers the ecological niche of a species as an *n*-dimensional hypervolume within which the populations of a species can persist. Habitat gradients and functional relationships among species (Whittaker et al. 1973) define this niche hypervolume. The niche concept led authors to promote several measurements for niche separation and niche breadth (Hurlbert 1978). For example, Colwell and Futuyma (1971) estimated niche breadth by measuring the uniformity of individual distribution among a set of resource states (ecological categories). Feinsinger et al. (1981) quantified niche breadth using a proportional similarity index. Such index measures the similarity between the frequency distribution of resources that are available and used by the individuals of a population.

Community studies usually focus on environmental gradients and refer to niches as differential habitat preferences of species (ter Braak and Verdonschot 1995). Such studies postulate that pairs of sampling units (SUs), which are similar in terms of physical and chem-

ical characteristics, should support similar species composition. In that context, observational approaches exploring the relationships between community structure and environmental gradients received much interest (e.g., Rodriguez and Magnan 1995). In addition, the ecological application of multivariate analyses increased as well. Ordination techniques are fundamentally well designed to investigate hyperspaces (Austin 1985, Grossman et al. 1991) and species–environment relationships (Dolédec and Chessel 1994, ter Braak 1994). For example, Green (1971, 1974) applied multiple discriminant analysis to identify ecological factors that separated the niches of mollusk species. The author used the standard deviation of scores on a discriminant axis as a measure of niche breadth and the extension of 50% probability ellipses as an index of niche overlap. Following this approach, Dueser and Shugart (1979) used 95% ellipse confidence to demonstrate the niche breadth of small mammals. The authors indicated that, for a given species, the average distance of the observations from the origin of the discriminant space was a reliable measure of its niche position relative to the average habitat conditions. Van Horne and Ford (1982) pointed out that the use of a discriminant space was not valid since the most abundant species were the most important to define this discriminant space. Finally, Carnes and Slade (1982) suggested computing the standard deviation of canonical scores as niche

breadth. The authors recommended calculating an index of niche specialization, i.e., the distance from the overall mean habitat, using equal weighting of sampling sites.

Among available multivariate techniques, correspondence analysis (CA; Hill 1974), also known as reciprocal averaging (RA; Hill 1973), is an ordination technique for investigating the separation of species niches or the ecological amplitude of species (Chessel et al. 1982, Thioulouse and Chessel 1992). From this analysis, ter Braak (1986) further developed a multivariate technique explicitly devoted to niche separation along environmental gradients known as canonical correspondence analysis (CCA). CCA was especially designed to extract the best combination of environmental variables (synthetic gradient) that maximizes the variance of the weighted average species positions ("niche centroids"). Pappas and Stoermer (1997) recently initiated a measure of multidimensional niche overlap based on species scores resulting from CCA. This multivariate direct gradient analysis received a widespread interest (see a review in Birks et al. 1996) despite its sensitivity to noisy environmental data (McCune 1997). Chessel et al. (1987), Lebreton et al. (1988), and ter Braak and Verdonshoot (1995) recognized the formal equivalence between CCA and discriminant analysis and emphasized that Green's analysis was a pioneering approach over ecological literature on community ordination.

CCA best suits the investigation of the unimodal species responses (Fig. 1a) to the environment (Palmer 1993). In addition, CCA implies that the importance of environmental measurements is proportional to the number of individuals per site. In other words, sites with more individuals will play a greater role. In contrast, redundancy analysis (RDA implemented in CANOCO; ter Braak 1987), also known as principal component analysis with respect to instrumental variables (Sabatier et al. 1989, Lebreton et al. 1991), assumes linear response to environment. RDA thus estimates the response curves of species by straight lines (Fig. 1b). Consequently the selection of RDA or CCA implies the selection of an either linear or unimodal model for the species response to the environment. In reality, however, community studies usually face a mixture of response models (Fig. 1c). Moreover, in case of limiting factor, the abundance patterns of species imply that the above traditional eigenanalyses hardly account for species-poor or individual-poor samples. As a result, there is a need for a method that addresses the various responses of the species to the environment and at the same time gives a fair vote to all sampling units.

In this paper, we start from the works of Perrin (1984) on terrestrial mollusks and Hausser (1995) on mammals. These authors investigated the geographical distribution of species by identifying which criteria cause a species to select a given habitat among
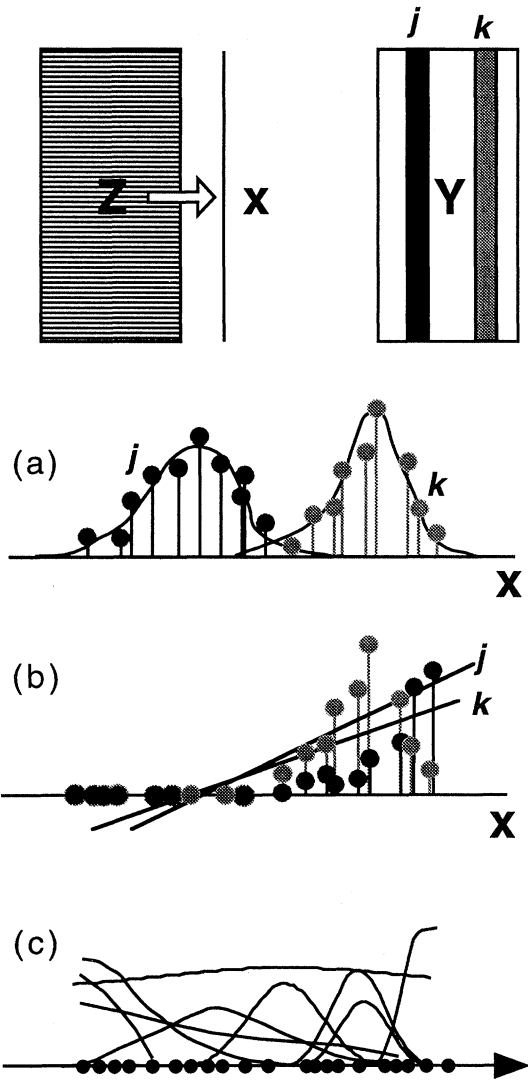

Fig. 1. Study of species–environment relationships. Matrix Z is the environmental data set. Matrix Y is the faunistic/floristic data set (species composition). The x-axes of graphs (a) and (b) represent synthetic gradients that combine several environmental variables. (a) The response curves of species j and k are unimodal, thus implying the use of canonical correspondence analysis (CCA). (b) The response curve of species j and k are close to a straight line because of a strong limiting factor at the lower end of the x gradient. This situation implies the use of redundancy analysis (RDA). (c) The typical conditions faced in gradient studies, with species having unimodal or only a part of this unimodal response curves (i.e., almost linear responses) included. This situation represents a typical case for using OMI analysis.

others, i.e., specialization criteria. We especially give a focus on a specialization criterion called marginality, which measures the distance between the average habitat conditions used by a species and the average habitat conditions of the sampling area (Hausser 1995). Here, we shift the emphasis from one species to species assemblages and we propose a new

multivariate technique explicitly based on the evaluation of the marginality of species that we call OMI (for Outlying Mean Index) analysis, or niche analysis for relating species assemblages to their environment. This new approach, which attempts to give a more even weight to all SUs even if they are species poor or individual poor, rather suits the diversity of the species responses to the environment. We compare the usefulness of the method to that of traditional eigenanalyses using a case of an ecological factor that favors the abundance of several species (e.g., meteorological conditions), and a case of a strong driving force for the abundance of several species (e.g., longitudinal stream gradient).

THE SAMPLING UNIT WEIGHT OPTIONS

Let us consider two tables: an environmental table **Z** and a faunistic table **Y**. The environmental table **Z** ($n \times p$) contains the measurements of p variables (as columns) in a set of n SUs (as rows). The species-composition table **Y** ($n \times t$) contains the abundance of t species (as columns) in the same SUs (as rows). Table **Y** can be analyzed by CA using the following notation:

$$y_{i\cdot} = \sum_{j=1}^{t} y_{ij} \qquad y_{\cdot j} = \sum_{i=1}^{n} y_{ij} \qquad y_{\cdot\cdot} = \sum_{j=1}^{t} y_{\cdot j}$$

which represents the row totals, the column totals and the overall total of the species-composition table respectively.

As a result, in CA, the SU and species weights are computed respectively as:

$$f_{i\cdot} = \frac{y_{i\cdot}}{y_{\cdot\cdot}} \qquad f_{\cdot j} = \frac{y_{\cdot j}}{y_{\cdot\cdot}} \qquad 1 \le i \le n,\ 1 \le j \le t.$$

Thereby, a traditional option is to weight the average position of species j by its relative abundance ($f_{\cdot j}$), i.e., the accuracy of the position of species j is proportional to its frequency. Moreover, the SU weighting ($f_{i\cdot}$) involves a greater weight for SU's having many abundant species. The study of table **Y** may require such weighting options in case of unequal sampling effort. Existing two-table ordination techniques such as CCA routinely weight SUs in proportion to their abundance vector. In that case, means and variances of the environmental variables (table **Z**) are usually computed using the SU weighting ($f_{i\cdot}$). Therefore, the selection of this weighting option implies that the importance of an environmental variable in a site appears to be proportional to the number of individuals collected at this site. As a result, traditional eigenanalyses using the above weighting option will scarcely include sites of limiting values (e.g., limited resource state, toxic pollution) since these sites have fewer species that include a fewer proportion of individuals.

MARGINALITY AND TOLERANCE

In the subsequent sections, we consider the analysis of the environmental table **Z** ($n \times p$) independently from that of the faunistic table **Y**.

Let the environmental table **Z** be analyzed by a PCA on a correlation matrix. Let $\mathbf{Z}_0$ ($n \times p$) contains the resulting standardized data (Fig. 2a). The center of gravity (G in Fig. 2a) of SUs is at the origin of the axes of this analysis and corresponds to the overall mean habitat. Let M_i represent SU i of table $\mathbf{Z}_0$ in the multidimensional space $\mathbb{R}^p$ (Fig. 2a). The total inertia of table $\mathbf{Z}_0$ equals

$$I_T = \sum_{i=1}^{n} p_i \|M_i\|_{\mathbf{I}_p}^2 \tag{1}$$

with p_i being the weight of SU i.

Let the faunistic table **Y** ($n \times t$) be transformed into a species profile table (noted as **F** in Fig. 2a) that contains the species proportions among SUs

$$f_{i/j} = \frac{y_{ij}}{y_{\cdot j}}$$

with y_{ij} as the abundance of species j in SU i and $y_{\cdot j}$ as the column total of species j. The inertia of species j is represented by the equation

$$I_T(j) = \sum_{i=1}^{n} f_{i/j} \|M_i\|_{\mathbf{I}_p}^2. \tag{2}$$

This inertia represents the total inertia of the table $\mathbf{Z}_0$ weighted by the profile of species j (Fig. 2b). Consequently, SUs that do not contain species j do not contribute to the inertia of species j.

Let us consider now an $\mathbf{I}_p$-normed vector **u** ($\|\mathbf{u}\|_{\mathbf{I}_p}^2 = 1$). The operation of projection of the rows of table $\mathbf{Z}_0$ onto vector **u** results in a vector of coordinates $\mathbf{Z}_0\mathbf{u}$.

Consequently, the average position of species j on **u** (or center of gravity of species j) is defined as

$$T_j = \mathbf{f}^\mathsf{T} \mathbf{Z}_0 \mathbf{u} \qquad \mathbf{f}^\mathsf{T} = (f_{1/j}, \ldots, f_{i/j}, \ldots, f_{n/j}). \tag{3}$$

From Eq. 3, the marginality or outlying mean index of species j [noted $m_a(j)$] along **u** is calculated by

$$m_a(j) = T_j^2 = (\mathbf{f} \mid \mathbf{Z}_0\mathbf{u})_{\mathbf{I}_p}^2 = (\mathbf{Z}_0^\mathsf{T}\mathbf{u} \mid \mathbf{f})_{\mathbf{I}_p}^2. \tag{4}$$

This marginality represents the deviation of the average position of species j from the origin (G). This is, thus, a measure of the distance between the average habitat conditions used by species j and the average habitat conditions of the sampling area.

From Eq. 4, the solution for **u** that maximizes $m_a(j)$ is

$$\mathbf{u}_j = \frac{\mathbf{Z}_0^\mathsf{T}\mathbf{f}}{\|\mathbf{Z}_0^\mathsf{T}\mathbf{f}\|_{\mathbf{I}_p}}. \tag{5}$$

Vector $\mathbf{u}_j$, thus, defines the direction (marginality axis) of species j represented by a thick line in Fig. 2a) for which the average position of species j (noted G_j in Fig. 2a) is as far as possible from the origin.

The dispersion of SUs that contain species j can be addressed as well. Let m_i be the projection of M_i onto the marginality axis. We denote the tolerance $T_m(j)$ of species j by

$$T_{\mathrm{m}}(j) = \sum_{i=1}^{n} f_{i/j} \|G_j - m_i\|_{I_p}^2. \tag{6}$$

which represents a measurement of the niche breadth of species j associated with the environmental variables of $\mathbf{Z}_0$.

In addition, a subspace is defined orthogonally to the marginality axis. The projection of SUs on the plane orthogonal to the marginality axis (Fig. 2d) yields a residual tolerance noted $T_r(j)$. This residual tolerance represents the variance in the species niche that is not taken into account by the marginality axis.

Finally, the application of the Pythagorean theorem with Eqs. 2, 4, and 6 allows the following decomposition:

$$I_{\mathrm{T}}(j) = m_{\mathrm{a}}(j) + T_{\mathrm{m}}(j) + T_r(j). \tag{7}$$

As a consequence, the variability of the niche of species j is decomposed into three components: (1) an index of marginality, i.e., the average distance of species j to the uniform distribution; (2) an index of tolerance or niche breadth; and (3) a residual tolerance, i.e., an index that helps to determine the reliability of a set of environmental conditions for the definition of the niche of species j. This decomposition for one species is central in the work of Hausser (1995). We hereunder expand these findings to the multispecific case.

PRINCIPLE OF THE OMI ANALYSIS

Basically, OMI analysis searches for a middle course among the marginality axes of each individual species in the assemblage. This eigenanalysis is thus explicitly designed to take into account the above niche components (defined for one species) in the case of a species assemblage. Eigenanalyses consist in the generalized singular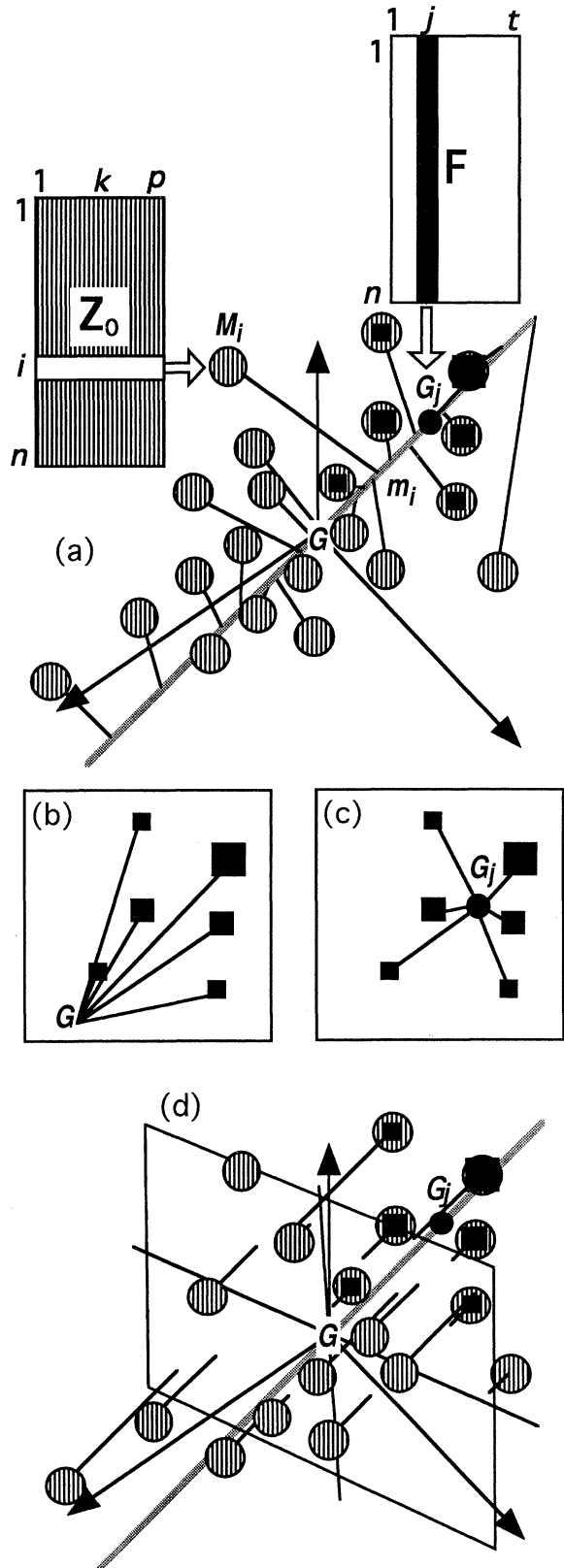 value decomposition (GSVD) of a statistical triplet $(\mathbf{X}, \mathbf{Q}, \mathbf{D})$ which summarizes the structure

$\rightarrow$

FIG. 2. Marginality and tolerance of a species. Matrix $\mathbf{Z}_0$ (n samples and p variables) is the environmental data set. Matrix $\mathbf{F}$ (n SUs [sampling units] and t species) is the species profile table (proportion per column). In (a) and (d) arrows represent the canonical basis. In (a), (b), and (d) G is at the center of gravity of the space defined by the rows of $\mathbf{Z}_0$. The origin of axes thus corresponds to the most general habitat conditions covered by SUs. (a) General definition and geometric presentation of the marginality and tolerance of a species. Gray circles represent the positions of SUs as linear combinations of environmental variables (ordination diagram). Square size is proportional to the frequency of species j among SUs where it occurs. G_j is the weighted average position of species j. The thick line represents the marginality axis. The thin lines stand for the projection of the rows of $\mathbf{Z}_0$ onto the marginality axis (e.g., m_i is the projection of M_i). (b) Inertia of species j as the weighted sum of squared distances to the origin (G). (c) Definition of the center of gravity of species j in the space of SUs positioned by environmental variables. (d) Definition of the residual tolerance of species j using the projection of the SUs onto a subspace orthogonal to the marginality axis.

TABLE 1. The various inputs for $(\mathbf{X}, \mathbf{Q}, \mathbf{D})$ statistical triplets, which are used in traditional eigenanalyses vs. OMI analysis.

Type	$\mathbf{X}$	$\mathbf{Q}$	$\mathbf{D}$
One-table techniques			
PCA on correlation	$\mathbf{Z}_0$†	$\mathbf{I}_p$	$(1/n)\mathbf{I}_n$
PCA on covariance	$\mathbf{Y}_0$‡	$\mathbf{I}_p$	$(1/n)\mathbf{I}_n$
CA	$\mathbf{D}_n^{-1}\mathbf{P}\mathbf{D}_t^{-1}-\mathbf{1}_{nt}$§	$\mathbf{D}_t$	$\mathbf{D}_n$
Row-weighted PCA	$\mathbf{Z}_0$‖	$\mathbf{I}_p$	$\mathbf{D}_n$
Two-table techniques			
RDA	$\dfrac{1}{n}\mathbf{Y}_0^{\mathsf{T}}\mathbf{Z}_0$	$\left(\dfrac{1}{n}\mathbf{Z}_0^{\mathsf{T}}\mathbf{Z}_0\right)^{-1}$	$\mathbf{I}_p$
CCA	$\mathbf{D}_t^{-1}\mathbf{P}^{\mathsf{T}}\mathbf{Z}_0$¶	$\mathbf{Z}_0^{\mathsf{T}}\mathbf{D}_n\mathbf{Z}_0)^{-1}$	$\mathbf{D}_t$
OMI analysis	$\mathbf{D}_t^{-1}\mathbf{P}^{\mathsf{T}}\mathbf{Z}_0$	$\mathbf{I}_p$	$\mathbf{D}_t$

Notes: Matrix $\mathbf{Z}$ contains raw environmental data, and matrix $\mathbf{Y}$ contains raw faunistic data. In canonical correspondence analysis (CCA) and OMI analysis, the environmental table is analyzed by a principal component analysis on a correlation matrix (PCA on correlation), and the faunistic table is analyzed by a correspondence analysis (CA). In redundancy analysis (RDA), the environmental table is analyzed by a PCA on a correlation matrix, and the faunistic table by a PCA on a covariance matrix (PCA on covariance). $\mathbf{I}$ represents the identity matrix.

† $\mathbf{Z}_0 = [z_{ij}^0 = z_{ij} - \bar{z}_j/\sqrt{\mathrm{var}\ z_j}]_{1 \le i \le n, 1 \le j \le p}$; $\bar{z}_j$ and $\mathrm{var}\ z_j$ are computed using $1/n$.

‡ $\mathbf{Y}_0 = [y_{ij}^0 = y_{ij}/y..]_{1 \le i \le s, 1 \le j \le r}$.

§ $\mathbf{P} = [f_{ij} = y_{ij}/y..]_{1 \le i \le s, 1 \le j \le r}$, $\mathbf{D}_n = Diag[f_{i.}]_{1 \le i \le n}$, and $\mathbf{D}_t = Diag[f_{.j}]_{1 \le j \le t}$.

‖ $\mathbf{Z}_0 = [z_{ij}^0 = z_{ij} - \bar{z}_j/\sqrt{\mathrm{var}\ z_j}]_{1 \le i \le n, 1 \le j \le p}$; $\bar{z}_j$ and $\mathrm{var}\ z_j$ are computed using $\mathbf{D}_n$.

¶ $\mathbf{Z}_0$ is analyzed first by a row-weighted PCA.

of matrix $\mathbf{X}$ using a metric given by matrix $\mathbf{Q}$ and row weights given by matrix $\mathbf{D}$ (see *Appendix*). In the general model of linear ordination methods, the various types of multivariate analyses, thus, differ by three parameters: (1) the matrix (noted $\mathbf{X}$ in Table 1), the centering options of which depends on the raw data (e.g., Noy-Meir 1973), (2) the metric (noted $\mathbf{Q}$ in Table 1), and (3) the row weights (noted $\mathbf{D}$ in Table 1). In the following, we focus on the row (SUs) weights.

In PCA and RDA, the rows of matrix $\mathbf{X}$ are uniformly weighted, whereas in CA and CCA they are weighted by their expected abundance $(f_{i.})$. In OMI analysis, matrix $\mathbf{X}$,

$$\mathbf{X} = \mathbf{D}_t^{-1}\mathbf{P}^{\mathsf{T}}\mathbf{Z}_0 \qquad \mathbf{P} = [f_{ij}]_{1 \le i \le n, 1 \le j \le t} \qquad (8)$$

which has t species (as rows) and p environmental variables (as columns), contains the centers of gravity of each species. The diagonal matrix of the species weights is denoted $\mathbf{D}_t = Diag(f_{.1}, \ldots, f_{.j}, \ldots, f_{.t})$ and the metric $(\mathbf{I}_p)$ is imposed by the analysis of the environmental table $\mathbf{Z}$ (e.g., PCA on a correlation matrix, see Table 1).

Consequently, we define OMI analysis as the GSVD of $(\mathbf{X}, \mathbf{I}_p, \mathbf{D}_t)$. The total inertia computed via OMI analysis is represented by

$$\mathrm{Tr}(\mathbf{X}^{\mathsf{T}}\mathbf{D}_t\mathbf{X}\mathbf{I}_p) = \sum_{j=1}^{t} f_{.j} m_a(j). \qquad (9)$$

This total inertia is, thus, proportional to the average marginality of species and represents a quantification of the influence of the environmental variables on the niche separation of species.

Total inertia is a useful statistic for testing the independence between two data tables (Kazi-Aoual et al.

1995). Accordingly, we used it to test the statistical significance of the species marginality. In practice, we used a Monte Carlo version of the test (Manly 1991), which considers the equiprobability of the $n!$ permutations of the rows (SUs) of the species profile table $\mathbf{F} = \mathbf{P}\mathbf{D}_t^{-1}$. Therefore, we first compared the observed marginality of a species (Eq. 4) to the distribution of the 1000 random permutation values obtained under the null hypothesis that the species is indifferent to its environment. We then focused on the total inertia of the OMI analysis (Eq. 9) and evaluated the statistical significance of the observed average marginality using the same approach. If, for example one of the 1000 permutation values was higher than the observed one, the marginality had an estimated significance of $P = 0.001$.

Note that all the above equations are given for the simplest case, i.e., quantitative environmental variables. However, the approach still holds for other kinds of environmental variables such as dummy (e.g., Tenenhaus and Young 1985), fuzzy coded (e.g., Chevenet et al. 1994), or a mixture of quantitative and dummy variables (Hill and Smith 1976).

ECOLOGICAL APPLICATIONS

To illustrate the potential of the OMI analysis we used two data sets that address the question of niche separation according to temporal or spatial characteristics of the habitat.

Temporal niche separation

The first data set investigated the influence of meteorological conditions on adult behavior using light trap catches of caddisflies (Usseglio-Polatera and Auda

TABLE 2. Number of individuals and percentage of species (and codes used as labels in the figures) of caddisflies caught in light traps.

Code	Species	No. individuals	Percentage of total
CHE	*Cheumatopsyche lepida* (Pictet)	3	0.04
HYC	*Hydropsyche contubernalis* McLachlan	3	0.04
HYM	*Hydropsyche modesta* Navas	183	2.35
HYS	*Hydropsyche siltalai* Döhler	6	0.08
PSY	*Psychomyia pusilla* (Fabricius)	5988	76.99
AGA	*Agapetus laniger* (Pictet)	109	1.40
GLO	*Glossosoma boltoni* Curtis	14	0.18
ATH	*Athripsodes albifrons* (L.)	14	0.18
CEA	*Ceraclea alboguttata* (Hagen)	7	0.09
CED	*Ceraclea dissimilis* (Stephens)	116	1.49
SET	*Setodes punctatus* (Fabricius)	189	2.43
ALL	*Allotrichia pallicornis* (Eaton)	52	0.67
HAN	*Hydroptila angulata* Mosely	191	2.46
HFO	*Hydroptila forcipata* (Eaton)	133	1.71
HSP	*Hydroptila sparsa* (Curtis)	470	6.04
HVE	*Hydroptila vectis* Curtis	9	0.12
STA	*Stactobiella risi* Felber	291	3.74

Note: Each night, light trapping lasted from one hour before sundown to one hour after daybreak.

1987). In this study, the authors intended to determine the reliability of daily variations in the species lists collected by light traps since such variations are generally attributed to the fluctuations of meteorological conditions. A total of 17 caddisfly species (see code in Table 2) were sampled at 49 occasions (nights) in June and July 1959 and 1960 at a site situated on the Rhône river (Lyon, France). At the same time, 11 meteorological variables were recorded (describing the temperature, intensity and direction of wind, pressure, precipitation, humidity, and cloudiness; see code in Table 3). A full description of these data is available in Usseglio-Polatera and Auda (1987: 72). Data consisted of the number of specimens of each species collected at each night. We transformed the species abundance into $\ln(x + 1)$ to reduce the effect of dominant species. We performed PCA on the correlation matrix of meteorological variables.

We computed and tested niche parameters (Table 4) to describe marginality and tolerance and, thus, the variability of responses of these caddisfly species to environmental variables. For example, *Hydropsyche modesta* (HYM in Table 4) had a low marginality and a high residual tolerance to meteorological conditions, whereas *Ceraclea alboguttata* (CEA in Table 4) was greatly influenced by meteorological conditions.

Ten out of 17 species showed a significant deviation of their niche from the origin suggesting a significant influence of the meteorological conditions for most of the species. Not surprisingly, the global test on the average marginality of all species was also significant ($P < 0.001$) since no permutation yielded a higher

TABLE 3. Meteorological variables (and codes used as labels in the figures) recorded at each light trap occasion.

Codes	Variables	Min.	Max.
MAX TEMP†	Maximum temperature (°C)	18.1	33.8
TWI TEMP	Twilight temperature (°C)	14.1	30.0
MIN TEMP†	Minimum temperature (°C)	7.8	20.0
WIND	Index of wind intensity and direction (m/s)‡	−9	+6
PRES	Mean atmospheric pressure (kPa)	132	134
PRES DIF§	Difference of atmospheric pressure (kPa)	−8.3	1.3
HUMI	Twilight humidity (%)	42	93
NIGHT CLOU	Night cloudiness (%)	0	100
NIGHT PRECIP	Night precipitation (mm)	0	12.3
MEAN CLOU	Mean cloudiness (%)	2	97
TOTAL PRECIP	Total precipitation (mm)	0	23.4

Notes: Each night light trapping lasted from one hour before sundown to one hour after daybreak. Minimum (Min.) and maximum (Max.) values are indicated.

† For one night of light-trapping.

‡ Wind intensity was arbitrarily provided with sign (+) if wind blew from the South. Sign (−) corresponded to other wind directions.

§ Difference of the atmospheric pressure between two consecutive nights measured at 6 h GMT.

TABLE 4. Niche parameters of 17 caddisfly species.

Species	Inertia	OMI	Tol	RTol	*OMI*	*Tol*	*RTol*	Num
Che	6.43	2.77	1.02	2.64	*43.1*	*15.9*	*41.0*	473 (NS)
Hyc	11.91	4.45	2.39	5.08	*37.3*	*20.0*	*42.6*	248 (NS)
Hym	10.57	0.09	2.54	7.94	*0.9*	*24.0*	*75.1*	227 (NS)
Hys	7.62	0.63	0.73	6.26	*8.3*	*9.6*	*82.1*	872 (NS)
Psy	10.47	0.43	3.92	6.11	*4.1*	*37.5*	*58.4*	0
Aga	7.43	1.29	1.55	4.59	*17.4*	*20.9*	*61.8*	0
Glo	14.36	6.18	4.76	3.42	*43.0*	*33.1*	*23.8*	4
Ath	11.24	1.79	2.76	6.68	*16.0*	*24.6*	*59.4*	63 (NS)
Cea	18.71	12.24	4.18	2.29	*65.4*	*22.3*	*12.3*	6
Ced	11.79	0.87	3.24	7.67	*7.4*	*27.5*	*65.1*	0
Set	12.61	4.29	3.72	4.60	*34.0*	*29.5*	*36.5*	0
All	6.80	0.72	1.21	4.87	*10.6*	*17.8*	*71.6*	125 (NS)
Han	10.37	1.21	3.37	5.79	*11.6*	*32.5*	*55.9*	84 (NS)
Hfo	17.54	6.76	7.34	3.44	*38.5*	*41.9*	*19.6*	9
Hsp	13.98	2.90	5.62	5.45	*20.7*	*40.2*	*39.0*	0
Hve	12.25	4.59	3.52	4.14	*37.5*	*28.7*	*33.8*	14
Sta	9.39	0.59	2.52	6.28	*6.3*	*26.9*	*66.9*	2

Notes: Species are identified by their codes (see Table 2). The inertia, the outlying mean index (OMI), the tolerance index (Tol), and the residual tolerance index (RTol) were computed for each species. Values in italics represent the corresponding percentages of variability. The last column (Num) represents the number of random permutations (out of 1000) that yielded a higher value than the observed OMI (NS = not significant at an estimate, $P \geq 0.05$).

value than the observed one. This statistical significance justified the plot of the species positions on an ordination diagram. The two first axes of the OMI analysis accounted for 95% of the marginality (90% for the first axis). As a consequence, subsequent graphs use only these two axes.

The construction of the ordination diagrams could be decomposed into three stages (Fig. 3). We first computed canonical weights (i.e., sum of square values equals 1) of the environmental variables (Fig. 3a). We then positioned the SUs (nights) by their scores that were linear combinations using the weights of environmental variables. The position of SUs and the distribution of species enabled the calculation of a weighted average (center of gravity) and an amplitude (Fig. 3b) for each species. As mentioned above, OMI analysis maximized the sum of the squared distances of the average positions of species to the origin. Finally, for each species, we superimposed the distance of these positions to the origin on the SUs plane (arrows in Fig. 3c). Each species position was proportional to the marginality index of that species (noted OMI in Table 4).

In this example, the ordination of SUs associated with the OMI analysis and that resulting from the separate analysis of environmental variables demonstrated a high similarity (Fig. 3d). As a conclusion, temperature (left side of the diagram in Fig. 3a) was the main factor affecting the abundance of caddisflies trapped by light (left side of the diagram in Fig. 3c). This factor did not have the same impact on all species (see OMI values in Table 4) demonstrating differences in temporal niches due to specific sensitivity to meteorological conditions. These differences reflected most probably specific emergence rhythms and duration of the adult life stage. However, using OMI analysis we were able to detect a common pattern among caddisfly species responses to meteorological conditions.

We computed RDA and CCA on the same data set. RDA and CCA extracted 49.3% and 34.8% respectively of the total faunistic variability. The proportion of variance of the community structure explained by environmental variables along the first axis showed a decrease from OMI analysis to RDA and a decrease from RDA to CCA (Table 5). Moreover, we computed canonical coefficients and correlation coefficients associated with the first axis of each analysis. Canonical coefficients represent environmental variable loadings of the best linear combinations that explain the community structure. Correlation coefficients measure the correlation between the best linear combination of environmental variables resulting from a two-table tech-

$\rightarrow$

FIG. 3. OMI analysis of caddisfly light-trap catches. The first axis is horizontal; the second axis is vertical. (a) Canonical weights of environmental variables (see codes in Table 3). (b) Distribution of species on the first factorial plane of an OMI analysis. Crosses identify the position of SUs using the canonical weights of environmental variables. The size of each circle is proportional to the proportion of a species (see codes in Table 1) in the SUs where it was captured. Lines link the center of gravity of a given species profile (weighted average position of a species noted G_j in Fig. 2) to each SU having this species. (c) Superposition of the weighted average positions of species (arrow ends) and SUs (crosses). The latter positions are identical to those in (b). (d) Projection of the two first PCA axes of the environmental variables on the first two axes of the OMI analysis.
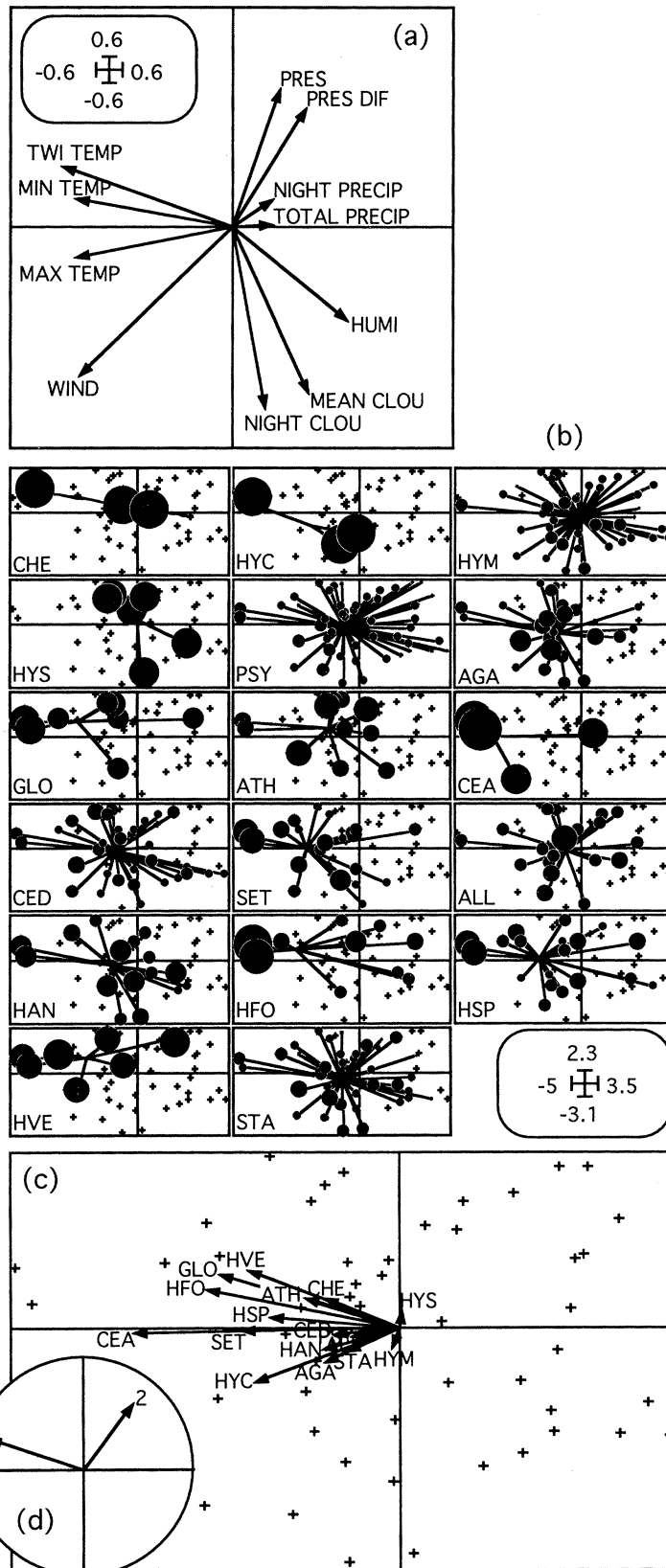
TABLE 5. Canonical coefficients (CC) and correlation (CO) of environmental variables (see Table 3 for codes) with the first axis of an OMI analysis (OMI), a redundancy analysis (RDA), and a canonical correspondence analysis (CCA) performed on the light-trap data set.

|	OMI		RDA		CCA	
Variables	CC	CO	CC	CO	CC	CO
MAX TEMP	−0.43	−0.91	−0.04	0.75	−0.52	0.63
TWI TEMP	−0.47	−0.95	−0.39	0.77	0.29	0.82
MIN TEMP	−0.43	−0.74	0.71	0.67	1.09	0.84
WIND	−0.42	−0.57	0.79	0.79	0.39	0.52
PRES	0.13	0.32	0.16	−0.30	0.45	0.01
PRES DIF	0.21	0.56	−0.14	−0.43	0.14	−0.12
HUMI	0.32	0.61	−0.26	−0.51	0.27	−0.57
NIGHT CLOU	0.09	0.25	−0.06	−0.07	0.22	−0.29
NIGHT PRECIP	0.12	0.30	−0.23	−0.23	−0.19	−0.11
MEAN CLOU	0.21	0.47	−0.20	−0.29	−0.47	−0.45
TOTAL PRECIP	0.12	0.33	0.16	−0.22	0.12	−0.13
Eigenvalues	1.27		3.90		0.14	
Percentage of variance	90.7		74.3		45.5	
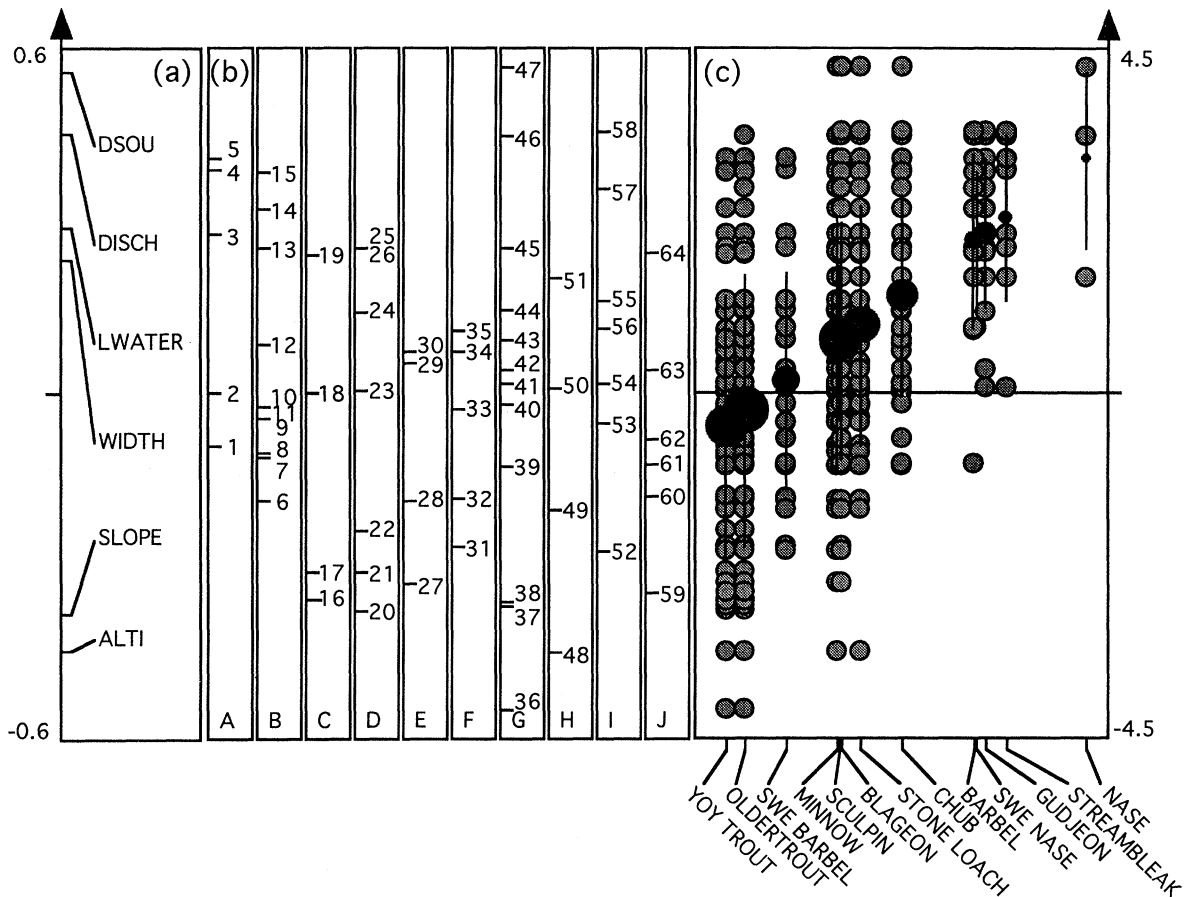
FIG. 4. OMI analysis of stream fish. We use here a one-axis presentation. (a) Canonical weights of environmental variables (DSOU = distance to the source, DISCH = mean annual discharge, LWATER = lowest monthly discharge occurring every five years, WIDTH = mean stream width, SLOPE = slope, and ALTI = altitude). (b) Site scores arranged by Rhône tributary (noted A–J; note that site labels per tributary increased from up- to downstream). (c) Species distribution arranged according to site scores (grey circles) in which a species occurred. The species order is given by their weighted average position along site scores. Species distributions are summarized by their means (sizes of black circles are proportional to the total frequencies of species) and standard deviations (vertical lines around the average corresponding to ±2 standard deviations).

TABLE 6. Occurrence (Occ) and niche parameters of 12 fish species.

Species	Occ	Inertia	OMI	Tol	RTol	Num
Sculpin (*Cottus gobio*)	6	5.50	39.0	27.3	33.6	71 (NS)
Older trout (*Salmo trutta*)	57	5.71	1.4	42.4	56.2	2
Y-O-Y† trout (*Salmo trutta*)	50	5.77	4.6	29.3	66.2	0
Minnow (*Phoxinus phoxinus*)	41	5.32	8.4	47.4	44.2	0
Stone loach (*Nemacheilus barbatulus*)	40	5.91	13.0	40.8	46.2	0
Blageon (*Telestes soufia*)	47	4.99	9.1	45.6	45.3	0
Southwestern barbel (*Barbus meridionalis*)	23	3.71	1.9	13.8	84.3	752 (NS)
Southwestern nase (*Chondrostoma toxostoma*)	4	6.38	76.5	11.7	11.7	22
Nase (*Chondrostoma nasus*)	3	12.60	81.0	12.9	6.2	5
Gudgjeon (*Gobio gobio*)	16	8.43	51.1	11.3	37.6	0
Chub (*Leuciscus cephalus*)	34	5.58	28.2	30.3	41.4	0
Streambleak (*Alburnoides bipunctatus*)	8	9.92	53.2	12.0	34.9	0
Barbel (*Barbus barbus*)	14	8.34	45.9	15.6	38.5	0

Note: The outlying mean index (OMI), the tolerance index (Tol), and the residual tolerance (RTol) are given as percentages of variability (see Table 4 for further details).
  † Young of the year.

nique and the best linear combination of these variables resulting from a one-table technique. The consistency (same sign and similarity of values) between canonical and correlation coefficients decreased from OMI analysis, to RDA and CCA (Table 5).

Spatial niche separation

The second data set aimed to investigate the fish assemblages along ten Mediterranean tributaries of the Rhône River (extracted from Pialot 1985). Twelve fish species (trout in two age groups) were collected and coded as present or absent at 64 sites distributed from upstream to downstream along these tributaries. Six variables describing the physical habitat along the longitudinal gradient of streams were simultaneously recorded. For each site, these were altitude, slope, yearly mean discharge (calculated from 15 yr), lowest monthly mean discharge at five-year reoccurrence, mean stream width, and distance to the source. We analyzed these environmental variables by PCA (after data standardization).

In this example, niche parameters identified the significance of marginality for most fish species (Table 6). This result corresponds to the view that longitudinal gradients are driving forces for stream organ-

isms (Hawkes 1975, Vannote et al. 1980, Statzner and Borchardt 1994). The most uniformly distributed species was the older trout but half the variability of the position of this species consisted of residual tolerance (Table 6). The deviation from the uniform distribution was maximal for nase and Southwestern nase, which were the most specialized species in this data set. The two first axes of the OMI analysis represented 97.9% of the explained variability and we selected the first axis (92.6%) for graphical presentations. This axis arranged the variables along the longitudinal gradient in a classical way (Fig. 4a). High altitude (abbreviated "Alti" in Fig. 4a) and slope defined the upper reaches in each stream (Fig. 4b). Width and discharge values increased with the distance downstream. The ordination of fish species (Fig. 4c) enabled the identification of three types of response to the longitudinal gradient: (1) trout were distributed over the entire longitudinal gradient; (2) blageon and associated fish species (minnow, sculpin, stone loach, and chub) were rarely found in the steep and high altitude reaches; and (3) barbel was associated with SWE nase, gudgeon, and streambleak, found together in downstream reaches.

We computed RDA and CCA on the same data set.

TABLE 7. Canonical coefficients (CC) and correlation (CO) of environmental variables with the first axis of an OMI analysis (OMI), an RDA, and a CCA performed on the fish data set.

Variables	OMI		RDA		CCA	
	CC	CO	CC	CO	CC	CO
ALTI	−0.45	−0.73	0.32	0.83	−0.44	−0.81
DSOU	0.56	0.96	−0.65	−0.95	0.33	0.93
SLOPE	−0.39	−0.80	0.18	0.72	−0.18	−0.72
WIDTH	0.23	0.67	0.24	−0.39	−0.28	0.27
DISCH	0.45	0.83	−0.05	−0.74	0.36	0.80
LWATER	0.29	0.59	−0.10	−0.49	−0.01	0.26
Eigenvalues	1.00		0.61		0.27	
% of variance	92.6		77.8		71.4	

Note: See legend of Fig. 4 for variable codes, and see Table 5 for further details.

RDA and CCA extracted 38.1% and 33.1% respectively of the total faunistic variability. The proportion of variance of the community structure explained by environmental variables along the first axis showed a decrease from OMI analysis, to RDA and CCA (Table 7). We again computed canonical coefficients and correlation coefficients associated to the first axis of each analysis. The consistency (same sign and similarity of values) between canonical coefficients and correlation coefficients decreased from OMI analysis, to RDA and CCA (Table 7).

DISCUSSION

Selection of an ordination technique

Should we add another sophisticated multivariate method to the existing ones? This old question was addressed by Noy-Meir and Whittaker (1977: 93) who underlined that "no single method has emerged as a solution to all problems of describing and explaining patterns of compositional variation in natural communities". Hence, there is no reason that only one method should investigate niche ordination with the same precision along a gradient or close to the threshold of a limiting factor.

The ordination of a faunistic/floristic table as a method to quantify patterns of covarying species abundance within species assemblages has three options: (1) the usual option uses species as variables and SUs are considered as frequency distributions among species; (2) species are considered as frequency distributions among SUs, i.e., species are the object under study; (3) both previous options are simultaneously addressed. The selection among these options requires the appropriate weighting of sites and species. A critical point is whether the sampling of the SUs is standardized (i.e., SUs can be related to a given area or duration) or not. In other words, the selection of an appropriate ordination technique should take into account whether the size of a sample is controlled or not. In the latter case, the effect of the unequal sampling effort has to be removed by CA or CCA. In that case, option (3) is implicit and elements of the matrix must be weighted by their expected abundance. This weighting option causes an SU supporting many abundant species to contribute more. Consequently, the importance of an environmental variable at a site is proportional to the number of species and the number of individuals at this site. In contrast, if the total sample abundance reflects real ecological differences, one would prefer the options 1 or 2. In that case, PCA on species profiles (e.g., Austin and Greigsmith 1968, Noy-Meir 1973) and OMI analysis, which focuses on option 2, should be used.

Comparison with other methods

One solution to avoid the problem associated with the SUs weighting option is to use RDA. In that case, a straight line estimates the bell-shaped response curves since RDA assumes a linear relationship between species and their environment. This analysis may be suited in the case of the existence of a strong limiting factor but may be unacceptable in the case of true unimodal species responses to the environment: a typical case for using CCA.

The advantage of OMI analysis in comparison to these ordination techniques is its good performance in describing either unimodal response curves or linear response curves that are linked to a limiting factor. It is rather difficult to know a priori the response curves of species for a given habitat gradient since, for example, the classical unimodal curve may be shaped due to interspecific competition. Consequently, in the absence of any explicit model describing the response curves of species under study, the niche analysis should include as few assumptions as possible (Austin 1985). Moreover, if the ordination method is not based on explicit response curves of species it should be as robust as possible to the differences in the underlying response curve models.

The meaning of OMI analysis

OMI analysis was designed to investigate and separate species niches according to their outlying mean index (average position on the marginality axis). It provides an integrated description of species–environment relationships as do other two-table ordination techniques. The advantage of using marginality (outlying mean index) is four-fold: (1) this index is a simple measure that integrates the niche specialization of species according to the selected habitat; (2) locations or resource states are positioned regardless of the presence of species, i.e., locations or resource states are weighted independently from their species composition (see Carnes and Slade 1982); (3) the reference represents neither the average nor the most abundant species but a theoretical ubiquitous species that tolerates the most general habitat conditions; and (4) the OMI of a species can be tested by random permutations thus assessing whether the observed species position in habitat differs significantly from what would be expected from chance alone. The null hypothesis thus stipulates that a species does not respond to the set of habitat variables under study. The significance of the global random permutation test of the OMI analysis makes it possible to demonstrate that niche segregation of species is effective along a given gradient. Finally, the total inertia of the OMI analysis (see Eq. 9) characterizes the global niche overlap of species. Therefore, this statistic could be used to compare the positions of species niches across various combinations of habitat conditions.

Variance terms are associated with the OMI to complete the position of a species according to niche differentiation. Tolerance describes the spatial or temporal variance of the niche across measured environmental
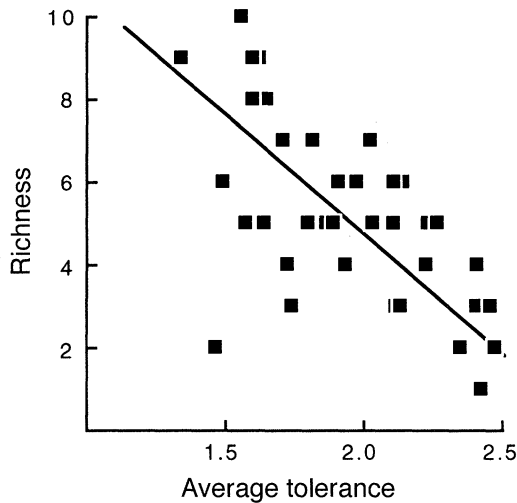
FIG. 5. Relation between the average tolerance values of sites and their fish species richness (Richness = 16.33 − 5.77[Tolerance]); SE [intercept] = 1.27, SE [coefficient] = 0.62; n = 64, R^2 = 0.59).

conditions or resources. In fact, marginality and tolerance of a species are correlated. In our fish example, we found a significant negative relationship ($P < 0.02$) between the marginality and the tolerance of species if we omitted nase, which was at the extreme of its distribution and behaved distinctly from other species (inertia = 12.60; see Table 6). This negative relationship agreed with the prediction of McNaughton and Wolf (1970) that more common species (low marginality) will have broader niches (high tolerance). We also tested the link between species richness and the size of species niches, which should be negatively related (McNaughton and Wolf 1970). In our fish example again, there was a significant negative relationship between the species richness of sites and their average tolerance value (Fig. 5; $P < 0.0001$). This relation can be interpreted in terms of increased competition among species having a broader niche (high tolerance) that decreases species diversity.

OMI analysis positions SUs in a multidimensional space as a function of environmental variables. This space should be close to the set of available resources for a species assemblage. The distribution of species along these conditions represents the realized niche, i.e., what a species really uses. OMI analysis allows a direct interpretation of ordination axes in terms of niche separation since the OMI represents the deviation of a species distribution from the overall mean habitat conditions and the method optimizes the separation of species centroids according to the average OMI of the species assemblage under study. The interpretation of OMI analysis can also consider niche specialization since a species close to the origin of the axes, i.e., most general habitat conditions covered by SUs, corresponds to the ubiquitous or generalist species. In contrast, spe-

cialists, which deviate from these general habitat conditions will demonstrate high OMI values. In our light-trap example, the marginality axis ordered species according to temperature, which affects the behavior of adult caddisflies. Furthermore, the analysis showed that the effect was not uniform for all species. In our fish example, species niches were separated along the longitudinal gradient demonstrating a continuum from low towards high ecological specialization. We also made groups of species being "guilds" in the sense of Simberloff and Dayan (1991: 115), i.e., groups of species "without regard to their taxonomic position, that overlap significantly in their niche requirements." These results indicate that a possible application of OMI analysis is the determination of niche shifts resulting from natural disturbances or from environmental management.

SOFTWARE AVAILABILITY

All calculations and graphs were made with ADE-4 (Thioulouse et al. 1995, 1997, Thioulouse and Chevenet 1996). Data of the examples that were treated in this paper are available in the package.

LITERATURE CITED

Austin, M. P. 1985. Continuum concept, ordination methods and niche theory. Annual Review of Ecology and Systematics 16:39–61.

Austin, M. P., and P. Greig-Smith 1968. The application of quantitative methods to vegetation survey. II Some methodological problems of data from rain forest. Journal of Ecology 56:827–844.

Begon, M., J. L. Harper, and C. R. Townsend. 1996. Ecology: individuals, populations and communities. Blackwell Scientific Publications, London, UK.

Birks, H. J. B, S. M. Peglar, and H. A. Austin. 1996. An annotated bibliography of canonical correspondence analysis and related constrained ordination methods 1986–1993. Abstracta Botanica 20:17–36.

Carnes, B. A., and N. A. Slade. 1982. Some comments on niche analysis in canonical space. Ecology 63:888–893.

Chessel, D., J. D. Lebreton, and R. Prodon. 1982. Mesures symétriques d'amplitude d'habitat et de diversité intra-échantillon dans un tableau espèces-relevés: cas d'un gradient simple. Compte Rendu de l'Académie des Sciences de Paris D 295:83–88.

Chessel, D., J. D. Lebreton, and N. Yoccoz. 1987. Propriétés de l'analyse canonique des correspondances. Une utilisation en hydrobiologie. Revue de Statistique Appliquée 35: 55–72.

Chevenet, F., S. Dolédec, and D. Chessel. 1994. A fuzzy coding approach for the analysis of long-term ecological data. Freshwater Biology 31:295–309.

Colwell, R. K., and D. J. Futuyma. 1971. On the measurement of niche breadth and overlap. Ecology 52:567–576.

Dolédec, S., and D. Chessel. 1994. Co-inertia analysis: an alternative method for studying species–environment relationships. Freshwater Biology 31:277–294.

Dueser, R. D., and H. H. Shugart, Jr. 1979. Microhabitats in a forest-floor small-mammal fauna. Ecology 59:89–98.

Escoufier, Y. 1987. The duality diagram: a means of better

practical applications. Pages 139–156 *in* P. Legendre and L. Legendre, editors. Development in numerical ecology, NATO Advanced Institute, Series G, Springer Verlag, Berlin, Germany.

Feinsinger, P., E. E. Spears, and R. W. Poole. 1981. A simple measure of niche breadth. Ecology **62**:27–32.

Green, R. H. 1971. A multivariate statistical approach to the Hutchinsonian niche: bivalve molluscs of Central Canada. Ecology **52**:543–556.

Green, R. H. 1974. Multivariate niche analysis with temporally varying environmental factors. Ecology **55**:73–83.

Greenacre, M. J. 1984. Theory and applications of correspondence analysis. Academic Press, London, UK.

Grossman, D. G., D. M. Nickerson, and M. C. Freeman. 1991. Principal component analyses of assemblage structure data: utility of tests based on eigenvalues. Ecology **72**:341–347.

Hausser, J. 1995. Säugetiere der Schweiz. Mammifères de la Suisse. Mammiferi della Svizzera. Birkhäuser Verlag, Berlin, Germany.

Hawkes, H. A. 1975. River zonation and classification. Pages 312–374 *in* B. A. Whitton, editor. River Ecology, University of California Press, Berkeley, California, USA.

Hill, M. O. 1973. Reciprocal averaging: an eigenvector method of ordination. Journal of Ecology **61**:237–249.

Hill, M. O. 1974. Correspondence analysis: a neglected multivariate method. Journal of the Royal Statistical Society (London) **23**:340–354.

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

Hurlbert, S. H. 1978. The measurement of niche overlap and some relatives. Ecology **59**:67–77.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology **22**:415–427.

Kazi-Aoual, F., S. Hitier, R. Sabatier, and J. D. Lebreton. 1995. Refined approximations to permutation tests for multivariate inference. Computational Statistics and Data Analysis **20**:643–656.

Lebreton, J. D., D. Chessel, R. Prodon, and N. Yoccoz. 1988. L'analyse des relations espèces-milieu par l'analyse canonique des correspondances. I. Variables de milieu quantitatives. Acta Œcologica, Œcologia Generalis **9**:53–67.

Lebreton, J. D., R. Sabatier, G. Banco, and A. M. Bacou. 1991. Principal component and correspondence analyses with respect to instrumental variables: an overview of their role in studies of structure–activity and species–environment relationships. Pages 85–114 *in* J. Devillers and W. Karcher, editors. Applied multivariate analysis in SAR and environmental studies, Kluwer Academic, Dordrecht, The Netherlands.

MacNaughton, S. J., and L. L. Wolf. 1970. Dominance and niche in ecological systems. Science **167**:131–139.

McCune, B. 1997. Influence of noisy environmental data on canonical correspondence analysis. Ecology **78**:2617–2623.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman & Hall, London, UK.

Noy-Meir, I. 1973. Data transformations in ecological ordination. I. Some advantages of non-centering. Journal of Ecology **61**:329–341.

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

Palmer, M. 1993. Putting things in even better order: the advantages of canonical correspondence analysis. Ecology **74**:2215–2230.

Pappas, J. F., and E. F. Stoermer. 1997. Multivariate measure of niche overlap using canonical correspondence analysis. Ecoscience **4**:240–245.

Perrin, N. 1984. Contribution à l'écologie du genre Cepaea

(Gastropoda): approche descriptive et expérimentale de l'habitat et de la niche écologique. Thesis. University of Lausanne, Lausanne, Switzerland.

Pialot, D. 1985. Analyse des données de milieu en hydrobiologie. Apport des techniques d'analyses multivariées. Thesis. University of Lyon 1, Lyon, France.

Rodriguez, M. A., and P. Magnan. 1995. Application of muiltivariate analyses in studies of the organization and structure of fish and invertebrate communities. Aquatic Sciences **57**:199–216.

Sabatier, R., J. D. Lebreton, and J. D. Chessel. (1989). Principal component analysis with instrumental variables as a tool for modelling composition data. Pages 341–352 *in* R. Coppi and S. Bolasco, editors. Multiway data analysis, Elsevier Science Publishers, The Netherlands.

Simberloff, D., and T. Dayan. 1991. The guild concept and the structure of ecological communities. Annual Review of Ecology and Systematics **22**:115–143.

Statzner, B., and D. Borchardt. 1994. Longitudinal patterns and processes along streams: modelling ecological responses to physical gradients. Pages 113–140 *in* P. S. Giller, A. G. Hildrew, and D. G. Raffaelli, editors. Aquatic ecology: scale, pattern and process. Blackwell Scientific Publications, Oxford, UK.

Tenenhaus, M., and F. W. Young. 1985. An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data. Psychometrika **50**:91–119.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology **67**:1169–1179.

ter Braak, C. J. F. 1987. CANOCO—a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis and redundancy analysis. Software documentation (version 2.1). Technical report LWA-88–02, Groep Landbouwwiskeunde, Wageningen, The Netherlands.

ter Braak, C. J. F. 1994. Canonical community ordination. Part 1: Basic theory and linear methods. Ecoscience **1**:127–140.

ter Braak, C. J. F., and P. F. M. Verdonschot. 1995. Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquatic Sciences **57**:255–289.

Thioulouse, J., and D. Chessel. 1992. A method for reciprocal scaling of species tolerance and sample diversity. Ecology **73**:670–680.

Thioulouse, J., D. Chessel, S. Dolédec, and J. M. Olivier. 1997. ADE-4: a multivariate analysis and graphical display software. Statistics and Computing **7**:75–83.

Thioulouse, J., and F. Chevenet. 1996. NetMul, a World-Wide Web user interface for multivariate analysis software. Computational Statistics and Data Analysis **21**:369–372.

Thioulouse, J., S. Dolédec, D. Chessel, and J. M. Olivier. 1995. ADE Software: multivariate analysis and graphical display of environmental data. Pages 57–62 *in* G. Guariso and A. Rizzoli, editors. Software per l'ambiente, Patron editore, Bologne, Italy.

Usseglio-Polatera, P., and Y. Auda. 1987. Influence des facteurs météorologiques sur les résultats de piégeage lumineux. Annales de Limnologie **23**:65–79.

Van Horne, B., and R. G. Ford. 1982. Niche breadth calculation based on discriminant analysis. Ecology **63**:1172–1174.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences **37**:130–137.

Whittaker, R. H., S. A. Levin, and R. B. Root. 1973. Niche, habitat and ecotope. The American Naturalist **107**:321–338.

APPENDIX

The following general procedure brings together various types of eigenanalyses such as correspondence analysis, multiple correspondence analysis, fuzzy correspondence analysis, the various kinds of principal component analyses (noncentered, row centered, column centered, etc.), redundancy analysis, canonical correspondence analysis, co-inertia analysis, and OMI analysis.

Let $\mathbf{X}$ be a table having n rows and p columns being derived from a raw data table by any kind of transformation. Let $(\mathbf{X}, \mathbf{Q}, \mathbf{D})$ be the resulting statistical triplet or duality diagram (Escoufier 1987). The $(n \times n)$ matrix $\mathbf{Q}$ contains the inner product (in $\mathbb{R}^p$) associated with the columns of $\mathbf{X}$. The $(p \times p)$ matrix $\mathbf{D}$ contains the inner product (in $\mathbb{R}^n$) associated with the rows of $\mathbf{X}$.

The generalized singular value decomposition (GSVD, see Greenacre 1984:344–346) of $(\mathbf{X}, \mathbf{Q}, \mathbf{D})$ consists in finding a $\mathbf{Q}$-normed axis $\mathbf{u}_1$ (first principal axis) and a $\mathbf{D}$-normed component $\mathbf{v}_1$ (first principal component) so that the inner products,

$$(\mathbf{XQu}_1 \mid \mathbf{v}_1)_{\mathbf{D}} = \mathbf{u}_1^{\top}\mathbf{QX}^{\top}\mathbf{Dv}_1 = \mathbf{v}_1^{\top}\mathbf{DXQu}_1 = (\mathbf{X}^{\top}\mathbf{Dv}_1 \mid \mathbf{u}_1)_{\mathbf{Q}}$$

and such that the quadratic forms $Q(\mathbf{u}_1)$ and $S(\mathbf{v}_1)$,

$$Q(\mathbf{u}_1) = \|\mathbf{XQu}_1\|_{\mathbf{D}}^2 = \mathbf{u}_1^{\top}\mathbf{QX}^{\top}\mathbf{DXQu}_1$$

$$S(\mathbf{v}_1) = \|\mathbf{X}^{\top}\mathbf{Dv}_1\|_{\mathbf{Q}}^2 = \mathbf{v}_1^{\top}\mathbf{DXQX}^{\top}\mathbf{Dv}_1$$

are maximized under the constraints that $\|\mathbf{u}_1\|_{\mathbf{Q}}^2 = \mathbf{u}_1^{\top}\mathbf{Qu}_1 = \|\mathbf{v}_1\|_{\mathbf{D}}^2 = \mathbf{v}_1^{\top}\mathbf{Dv}_1 = 1$.

The achieved maximum of the above inner product is equal to the first singular value. The solution vectors $\mathbf{u}_1$ and $\mathbf{v}_1$ can also be obtained as the right-hand eigenvectors of $\mathbf{X}^{\top}\mathbf{DXQ}$ and $\mathbf{XQX}^{\top}\mathbf{D}$, respectively, and the maxima of $Q(\mathbf{u}_1)$ and $S(\mathbf{v}_1)$ are equal and given by the first eigenvalue of these matrices (which is the square of the above-mentioned first singular value). The rows of $\mathbf{X}$ can be $\mathbf{Q}$-projected on $\mathbf{u}_1$ and the columns of $\mathbf{X}$ can be $\mathbf{D}$-projected on $\mathbf{v}_1$ resulting in the first scores $\mathbf{x}_1$ and $\mathbf{y}_1$ as follows:

$$\mathbf{x}_1 = \mathbf{XQu}_1 \qquad \mathbf{y}_1 = \mathbf{X}^{\top}\mathbf{Dv}_1.$$

If r is the rank of table $\mathbf{X}$, then the second and further principal axes $(\mathbf{u}_2, \mathbf{u}_3. \ldots, \mathbf{u}_r)$ and the second and further principal components $(\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_r)$ maximize the same inner products and norms but are subjected to extra constraints of orthogonality, i.e., for all $s \neq t$ $(\mathbf{u}_s \mid \mathbf{u}_t)_{\mathbf{Q}} = (\mathbf{v}_s \mid \mathbf{v}_t)_{\mathbf{D}} = 0$.