

A phytosociological method for interpreting plankton data

J.-M. Fromentin¹, F. Ibanez¹, P. Legendre²

¹ Observatoire Océanologique, Station Zoologique, BP 28, F-06230 Villefranche-sur-Mer, France

² Département de sciences biologiques, Université de Montréal, CP 6128, Succursale A, Montréal (Québec), Canada H3C 3J7

ABSTRACT: Multivariate techniques allow data sets to be summarised, but are not designed to causally model ecological phenomena. The ambiguity of the interpretation results from the ambiguity of the scaling space. The phytosociological method proposed by Grandjouan (1982; Thèse d'état, Université Louis Pasteur, Strasbourg) is a general interpretive framework allowing separate consideration of 3 distinct reference spaces: a geographical space to visualise the spatio-temporal distribution of the descriptors, an ecological space showing the effects of the environmental variables on the biological descriptors, and a bio-ethological space showing the statistically meaningful species communities. It is thus possible to define the position of each biological category and each observation along the scale of variation of each environmental variable (hydrological position, HP) or each geographic coordinate (geographical position, GP), as well as its density relative to these scales (indicative value, IV). Using continuous zooplankton data collected along a transect across the Ligurian Sea front, principal component analysis and correspondence analysis were compared to this technique. The phytosociological method produced the most meaningful ecological results about the structure of this frontal ecosystem.

INTRODUCTION

Multivariate analysis is commonly used to describe marine ecological data. The aim of this study is not to criticise these methods or to propose yet another multivariate treatment, but to test the possible advantages of applying to marine ecology a method originally developed for phytosociological analysis (Grandjouan 1982a). The method is a general framework, based on concepts derived from ecological considerations, which incorporates classical classification and scaling techniques as its tools. This method addresses problems concerning the structural dependency of species and environmental factors, in a spatio-temporal reference frame. The idea is not to apply several techniques and look for possible explanations. Instead, each algorithm is intended to answer some specific ecological question. For instance: What are the species most affected by a change in a given environmental factor? What kind of species assemblages, and how many, are

present? What is the cause of a high level of similarity observed between the abundances of 2 species: is it due to their sharing the same type of biological behaviour or a common response to hydrological factors, or simply because their variations obey the same spatial gradient (e.g. the strong coastal-offshore gradient), or the same temporal structure (e.g. they have the same reproductive phase)?

The chief property of multivariate data analysis techniques is that they summarise the data. Clustering allows sites, times, or species to be combined into groups; clustering results may be represented by dendrograms in the case of hierarchical methods and by a partition when using a non-hierarchical technique. Scaling techniques (ordinations) reduce the dimensionality of the data, concentrating most of the variance into a few component axes which, in some methods, are linear combinations of the original variables. Many computer programs are available to perform principal component analysis (PCA), corre-

spondence analysis (CA), principal coordinate analysis, nonmetric multidimensional scaling, canonical variate (discriminant) analysis, etc.; these techniques have been described in the context of ecology by Legendre & Legendre (1983, 1984a) and by ter Braak (1987).

In principle, clustering and scaling techniques have to be considered as preliminary exploratory steps in the process of elaborating explanatory models or designing new experiments. But most often in practice, ecologists search through the graphical results to discover clues as to the underlying structure of the ecosystem they are studying. To make the scaling configuration easier to understand, some programs allow projection of supplementary external variables, or the spatio-temporal coordinates, in the ordination space. Constrained clustering in time or space (Legendre & Legendre 1984b, Legendre et al. 1985, Legendre 1987) or constrained scaling taking into account the correlations between species and environmental factors (ter Braak 1986, 1987) permits maximisation principles to be introduced in order to increase the ecological significance of the representations.

The problems of interpreting scaling structures are due to the ambiguity of the mathematical space. A space defined by the frequencies of organisms is not identical to that defined by environmental variables, and vice versa. Putting together variables of different types (e.g. species and environmental factors) makes it more difficult to clearly identify the main driving components of the space. CA allows species and observations to be projected in the same space, but what is the meaning of the graphical distance between a station and a species? And what is the meaning of the space occupied jointly by the species (finite space) and by the observations (infinite space)?

Ecological phenomena cannot be clearly described by putting all the information in a single space. Species are preferentially associated with particular intervals along the variation axes of the environmental factors, and with some locations or periods, and they have trophic, reproductive, and competitive behaviours. All these are relevant to ecologists, but they cannot be highlighted by a global multivariate treatment.

The phytosociological numerical methodology proposed by Grandjouan (1982a, b) and by Brisse & Grandjouan (1978, 1980a, b) consists of separately considering 3 different spaces: the geographical space describing the spatio-temporal intersections of the species distributions, the ecological space defined by the environmental variables (with reference to the ecological niche concept of Hutchinson 1957), and the biological space reflecting both biological and ethological interactions. These unambiguous spaces lead to descriptive representations which allow the results of

the analyses to be read directly and easily, without prior knowledge or subjective evaluation. Classical multivariate techniques can be used to produce these summary spaces.

An illustration of the method as applied to marine ecological data is presented here. Zooplankton species and multivariate hydrological monitoring data were collected along a transect crossing the Ligurian frontal zone in the Mediterranean. The complexities involved in interpreting the interactions between biological behaviours and hydrodynamics, using classical methods, have been discussed by several authors (Boucher 1984, 1988, Ibanez & Boucher 1987, Williams 1988, Williams & Conway 1988). The data used here have been analysed in a previous paper on seasonal frontal variability (Boucher et al. 1987); they seemed adequate for testing the improvement in interpretation that can be obtained in marine ecology from the proposed new type of data analysis. To include all possible numerical methods in this comparison would be an impossible task; thus, here only a comparison with the results obtained by applying ordinary PCA and CA to the different data sub-tables will be made, because these 2 analyses represent classes of more sophisticated techniques. PCA represents the class of quantitative methods imposing a linear model onto the relationships among variables; the Pearson correlation is the measure of association among species. CA is a method valid for presence-absence or quantitative data; it uses the χ^2 distance which, for species abundance data, is suitable for extracting nonlinear, single-peaked relationships of species with the environmental factors (ter Braak 1987).

DATA

The sampling campaign, crossing the Ligurian Sea front along a transect from Nice to Calvi 30 nautical miles offshore (Fig. 1), was conducted on March 3, 1982, as part of a cruise in connection with the research program PRETROPHOS (Villefranche-sur-Mer, France). Temperature, salinity, water density and fluorescence were measured at 20 s intervals by a hydrological probe constructed by L. Prieur (Laboratoire de Physique et Chimie Marine de Villefranche-sur-Mer). On March 4, vertical measurements were made with the same probe at 6 locations along the transect to obtain a picture of the profile structure. Fig. 2 is a schematic representation of the vertical water transport corresponding to an alternation of convergence and divergence movements forming successive cells and involving layers of various thickness. Horizontally, 3 main hydrological zones can be recognized in the northern Ligurian basin (Boucher et al. 1987):

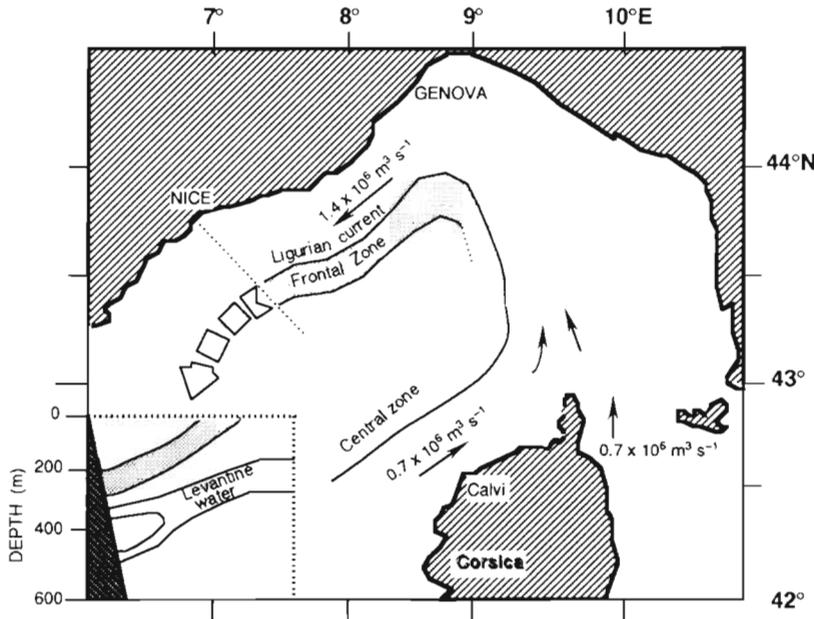


Fig. 1. Schematic map of the circulation in the Ligurian Sea. Vertical distribution of water masses for the studied area is schematised in the lower left part of the figure. The sampling transect crossing the front is represented by a dotted line (from Boucher et al. 1987)

(1) A coastal and peripheral zone influenced by the cyclonic Ligurian current and limited outward by a coastal convergence, C' . Throughout the year, the width of the peripheral zone varies from 10 to 20 km.

(2) A frontal zone characterized by the highest horizontal gradient of salinity and density, and bounded by an offshore divergence, $D1$, and by the main convergence, C . The width of the frontal zone is 1 to 20 km.

(3) A marginal zone, followed after the deep divergence ($D2$) by the central zone. It is defined by high surface density and weak stratification. The highest biomass of chlorophyll is found in the marginal part of the frontal zone, favoured by the upwelling of nutrients associated with the offshore divergence ($D1$). The width of the central zone, extending to about 130 km from Nice, is 60 to 110 km.

In order to couple the physical and biological data, the values observed during 5 min intervals were averaged. The spatial variations of these variables are presented in Fig. 3. For zooplankton, 68 samples were collected at 7 m depth by a Tube Hai system, at 5 min intervals; for each sample, 35 m^3 of water were filtered along 770 m, given a ship speed of 5 knots. Twenty-four zooplankton groups were counted (Table 1, Fig. 4; data provided by J.

Boucher, IFREMER, Nantes, whose contribution is thankfully acknowledged). The successive hydrological zones (peripheral, frontal and central) are represented in Figs. 3 and 4. During this cruise in March, 1982, these zones extended to 15, 35 and 40 km offshore. The marginal zone is not visible: the frontal zone is characterized by a progressive gradient of salinity and density, and the well-defined divergence $D2$ occurs just afterwards, with high water density and a vertical instability marked by the lowest values of fluorescence.

NUMERICAL ANALYSES

The phytosociological approach was compared to the following well-known techniques of classification and ordination.

Two PCAs were performed: one from the correlation matrix among species abundances (in logs), in which temperature, salinity, density and fluorescence were added as supplementary variables to the graph of descriptors, and another from the correlation matrix among the stations. CA was performed on a binary

temperature, salinity, density and fluorescence were added as supplementary variables to the graph of descriptors, and another from the correlation matrix among the stations. CA was performed on a binary

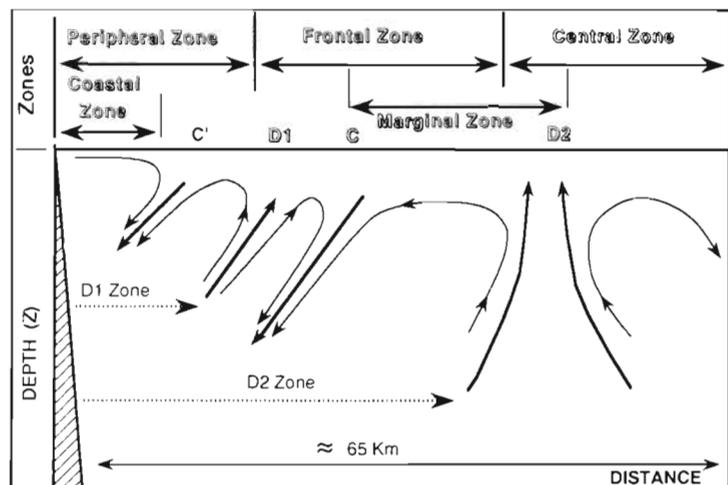


Fig. 2. Cross-frontal circulation showing the divergence and convergence lines (thick) and associated stream lines (thin). The names of the zones are indicated above (from Boucher et al. 1987). The peripheral zone stretches between the coast and the beginning of the horizontal density gradient; the frontal and central zones are further offshore. Key points C' , $D1$, C , $D2$ (downwellings and upwellings) indicate the main discontinuities demarcating the convection cells. The width of the zones thus created and that of the vertical instability regions vary during the course of the year

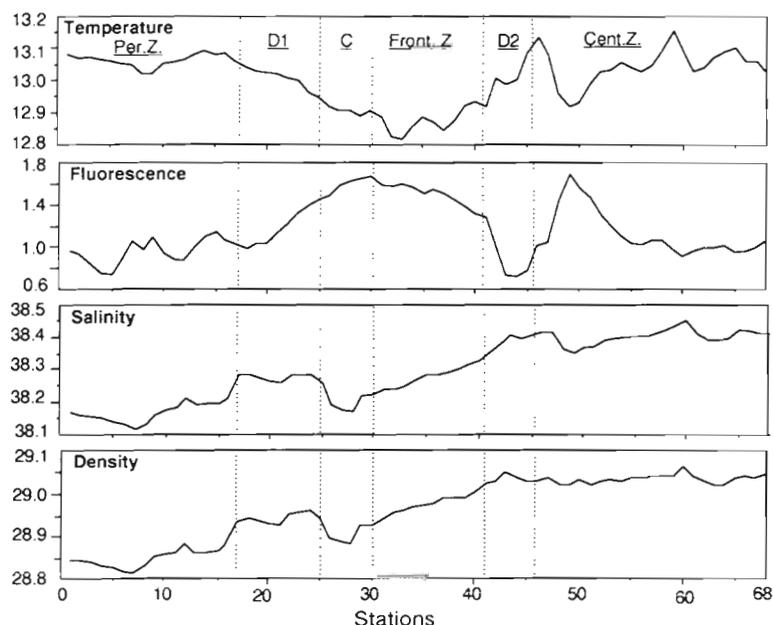


Fig. 3. Records of temperature ($^{\circ}\text{C}$), fluorescence (equivalent to chlorophyll a units: mg l^{-1}), salinity (PSU), and density ($\text{kg m}^{-3} - 1000$). Horizontal axis shows the position of the sampling stations, not kilometers. The hydrological compartments are delimited by dotted lines. Although divergence zone D1 (local increase in salinity and density), convergence zone C (decrease in salinity), and divergence D2 (highest density values) are well defined, the marginal zone is not. The frontal zone shows a continuously increasing trend in density

data table; rows correspond to the observations (68), while the columns (79) correspond to 67 species abundance classes and to 12 environmental variable classes where the observations were coded as either present (1) or absent (0). Species abundances were coded into 2 to 4 almost equal-frequency classes (the number of these classes depends on the abundance of the species), while environmental variables were each coded into 3 classes. Note that these classes are completely different from the phytosociological classes (see below), and don't overlap them.

A classification method was used to delineate groups of objects in the ordination space, for a chosen level of dissimilarity; among the hierarchical clustering strategies described by Lance & Williams (1967), the flexible clustering method with $\beta = -0.3$ was used on a matrix of Euclidean distances among observation vectors.

The phytosociological method described below requires the estimation of distance matrices in order to ordinate the observations and the species by nonmetric multidimensional scaling (MDS; Kruskal 1964).

Table 1. Mean abundances (no. m^{-3}), standard deviations, and number of abundance classes with their thresholds (generated by a segmentation of each distribution), for the 24 zooplankton taxonomic groups. *Clausocalanus* spp. are separated into 3 classes (A, B, C) of decreasing size (Boucher et al. 1987). Abbreviations used in figures are also shown

| Taxonomic group | Abbreviation | Mean | SD | No. of abundance classes | Thresholds |
|---------------------------------------|--------------|------|------|--------------------------|---------------|
| <i>Acartia clausi</i> | Aca | 249 | 240 | 2 | 181; 432 |
| <i>Calanus helgolandicus</i> adults | Cal | 105 | 152 | 3 | 0; 70; 247 |
| <i>C. helgolandicus</i> copepodites 1 | Cp1 | 40 | 40 | 2 | 36; 66 |
| <i>C. helgolandicus</i> copepodites 2 | Cp2 | 67 | 61 | 2 | 59; 108 |
| <i>C. helgolandicus</i> copepodites 3 | Cp3 | 57 | 56 | 2 | 53; 119 |
| <i>C. helgolandicus</i> copepodites 4 | Cp4 | 63 | 57 | 2 | 55; 124 |
| <i>C. helgolandicus</i> copepodites 5 | Cp5 | 112 | 125 | 3 | 0; 107; 306 |
| <i>Clausocalanus</i> spp. A | ClA | 427 | 465 | 2 | 350; 934 |
| <i>Clausocalanus</i> spp. B | ClB | 910 | 1455 | 2 | 947; 2970 |
| <i>Clausocalanus</i> spp. C | ClC | 1567 | 1781 | 2 | 1170; 3212 |
| <i>Centropages typicus</i> adults | Cea | 272 | 354 | 3 | 153; 356; 828 |
| <i>C. typicus</i> juveniles | Cej | 180 | 223 | 3 | 120; 268; 598 |
| Copepod nauplii | Nau | 132 | 103 | 2 | 125; 289 |
| <i>Oithona</i> sp. | Oit | 1546 | 1282 | 1 | 2165 |
| Acantharians | Acn | 16 | 12 | 2 | 0; 14 |
| Cladocerans | Clc | 6 | 25 | 3 | 0; 4; 11 |
| Echinoderm larvae | Ech | 8 | 20 | 2 | 0; 18 |
| Decapod larvae | Dec | 59 | 55 | 2 | 55; 126 |
| Gastropod larvae | Gas | 10 | 15 | 2 | 0; 23 |
| Crustacean eggs | Eggs | 91 | 103 | 2 | 91; 174 |
| Ostracods | Ost | 2 | 3 | 3 | 0; 3; 7 |
| Pteropods | Pte | 0.5 | 1.3 | 2 | 0; 3 |
| Siphonophores | Sip | 6 | 9 | 3 | 0; 7; 16 |
| Bells of calycophores | Bells | 8 | 9 | 2 | 0; 8 |

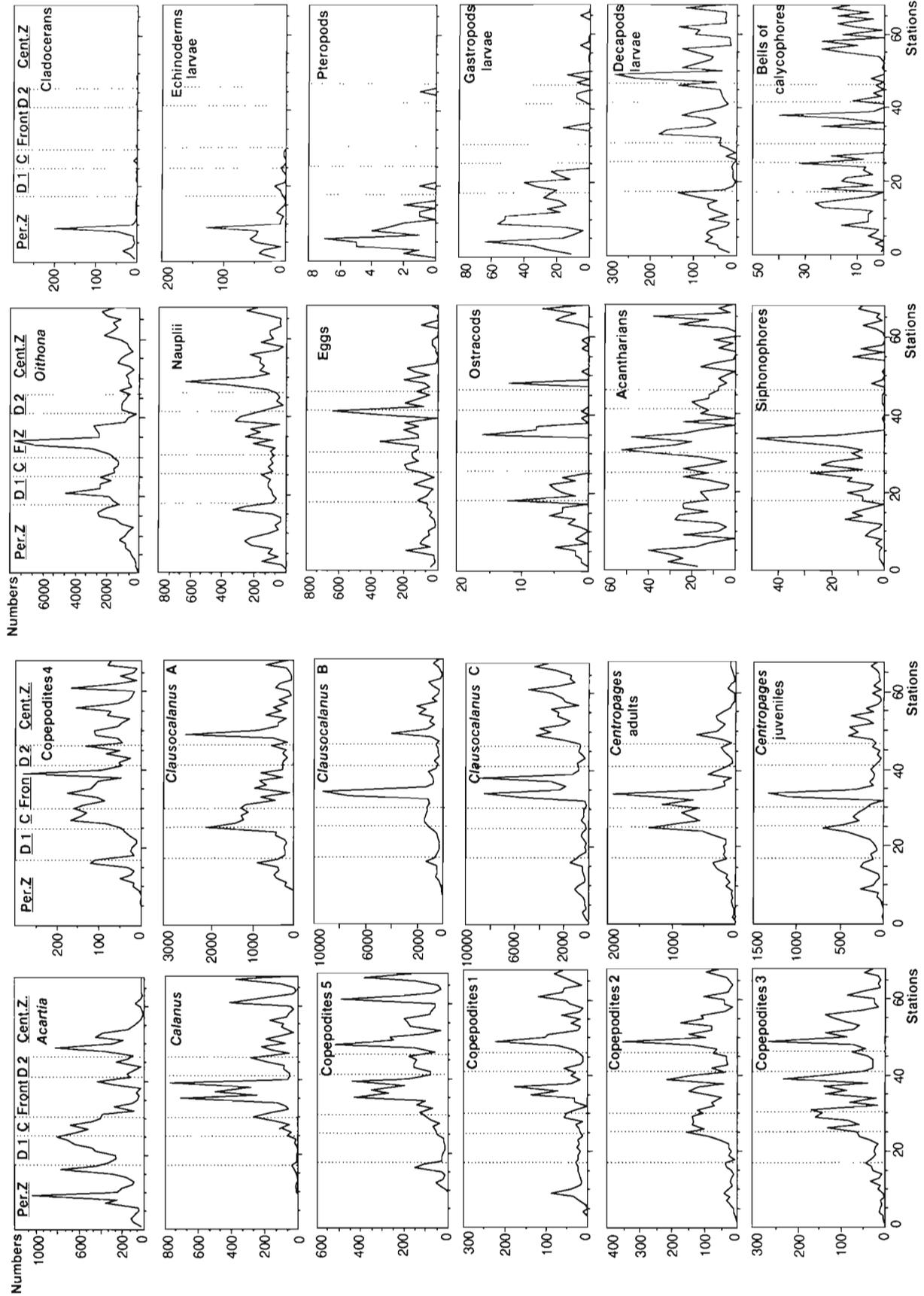


Fig. 4. Abundance (no. m⁻³) of the 24 taxonomic groups along the transect. The hydrological compartments are delimited by dotted lines (see legend of Fig. 3)

This type of scaling only attempts to preserve the *order* of the distances among objects (observations or species) in the reduced space, instead of their Euclidean distances.

To compare the distance matrices corresponding to the 3 spaces (geography, ecological factors, biology), one cannot use a classical correlation test, because of the relations that exist among distance measures within any distance matrix (axiom of the triangle inequality; Legendre & Legendre 1984b). Hence the Mantel randomisation test was used instead (Mantel 1967, Legendre & Fortin 1989). When the values within each of the distance matrices are standardised, the Mantel statistic is identical to a linear correlation between 2 vectors of distances. Repeatedly permuting at random the rows and corresponding columns of one of these matrices, followed by re-computation of the statistic, produces an empirical null distribution against which the actual value of the statistic can be tested; alternatively, when the number of observations in each distance matrix is large, a normal approximation proposed by Mantel (1967) can be computed instead of performing the permutation test. Since the distances in the 3 spaces are likely to be correlated, we computed partial Mantel tests (Smouse et al. 1986) in order to eliminate the linear relationships between distance spaces. In that method, residual distance matrices are computed by linear regression, as in classical partial correlation and regression, and the Mantel test is performed between residual distance matrices. The total and partial Mantel correlations so obtained allowed elaboration of a simple causal model linking the ecological structures embedded into the different spaces.

The tests of statistical significance of the Mantel statistics reported below have to be interpreted with caution, because the 3 distance matrices under study are not independent from one another; since they represent projections of the same taxonomic presence-absence data matrix in 3 different spaces, this is likely to artificially make the 3 distance matrices more similar to one another. This built-in correlation among spaces could not make a correlation lose its significance, however, so that if a Mantel correlation is found to be not significant, it cannot be the result of the way the matrices were built. It is precisely these non-significant Mantel correlations that we will be focusing on in the Results.

THE PHYTOSOCIOLOGICAL METHOD

The method proposed by Brisse & Grandjouan (1978, 1980a, b, 1981), Grandjouan (1982a, b), and Brisse et al. (1985) tries to avoid some of the inadequacies

associated with the direct use of classical ordination and classification methods in the ecological framework.

Linear correlations: Several data analysis methods use correlation coefficients; with species presence or abundance data, however, linear correlation coefficients are not adequate to model species relationships because coefficient values increase with the co-absence of species. When comparing species presence or abundance data to environmental variables, correlation coefficients are only capable of reflecting direct, linear relationships. If a species is rare or is restricted to one or a few intermediate classes of variation of the external variable, the correlation vanishes.

The χ^2 distance: This distance measure, used in CA and derived analyses, gives more importance to rare species (weighting is inversely proportional to the frequency), and thus underestimates the influence of the most abundant species (Legendre & Legendre 1984a).

The mean: Ecological data most often consist of discontinuous patches in space and/or time, and so generally, this statistic cannot be ecologically interpreted.

Transformation of abundance data: Generally before analysis, species abundance data are log-transformed in order to reduce heteroscedasticity (heterogeneity of the variances). This operation can be ecologically justified by our knowledge that populations grow exponentially (Malthus 1798), or by non-linear relations that may be found between species and environmental variables. Even after a log transformation, it may still be useful to divide species abundance data into density classes in order to clearly display the response of the species to environmental conditions; obtaining classes makes it possible to use these data together with qualitative multistate descriptors in the same analysis. To determine the segmentation thresholds, let us consider the retrocumulative distribution histogram (Fig. 5) of the densities of a species; 12 abundance classes were arbitrarily chosen at this step. Ordered abundance classes (levels) are empirically defined as follows (Brisse & Grandjouan 1982):

- If there are more than 10 % null values, a first class is attributed to absences. Samples with that species absent are coded 1, the others are coded 0.
- Reading the histogram from left to right, the first threshold is reached when a column j is found that presents a retrocumulative frequency smaller than half the frequency in any of the previous columns. The threshold is located at the lower (left) limit of column j . An abundance class (level) is thus created, and the observations are coded 0 for values lower than the threshold and 1 for higher values.
- This procedure is repeated, taking column j as the beginning of the histogram, and looking for new frequency thresholds.

Retrocumulative frequencies

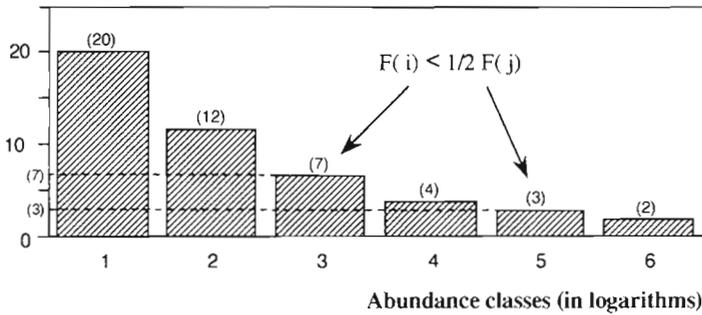


Fig. 5. Determination of thresholds for species abundance classes. From left to right, if the frequency of a column (j) of the histogram is lower than half the frequency of any of the previous columns (i), its lower (left) limit indicates an abundance threshold. Then a binary abundance class is created, and the observations are coded 0 for values lower than the threshold and 1 for higher values. The procedure is repeated, taking column j as the beginning of a new histogram. In this example, class 3 (frequency 7) is smaller than half the frequency of class 1 (20), so that the first threshold is established; a binary variable is created, coding 0 for stations in abundance classes 1 and 2, and 1 elsewhere. Taking now class 5 (frequency 3), the same procedure is carried out with respect to class 3; another binary variable is created, coding 0 for stations with abundance in classes 1 to 4, and 1 elsewhere

- Thus there is one abundance class descriptor less than the actual number of classes. Observations belonging to the last class receive code 1 for all the previous abundance class descriptors and are thus distinguishable. (For correspondence analysis, a separate abundance class descriptor is created for that last class, even though it is redundant with the previous classes.)

Notice the overlapping between these classes, because the observations coded as 1 in the last abundance classes are also coded as 1 in the previous ones. This procedure generally produces no more than 3 or 4 abundance classes, but in order to eliminate those with negligible frequencies, a minimum of 10 % species presence is required for a class to be retained, unless there are biological reasons to act differently; in the case of *Clausocalanus* spp. B and the *Centropages typicus* adults, for instance, this minimum was changed because of previous knowledge. For the 24 recorded taxonomic groups, 54 abundance classes were created; they are listed in Table 1. These binary species abundance variables were used to estimate the characteristics of the species in the ecological (defined by the values of the environmental variables, assumed to be quantitative) and the geographic space (defined by the spatio-temporal coordinates of the observations). The procedure for analysing the biological space is presented in a later section.

Ecological and geographic spaces

Characterisation of species by their fidelity (FD)

The 'fidelity' of a species to a class of a variable, ecological or geographic, corresponds to its frequency (number of presences) in the stations coded with this ecological or geographic characteristic present, divided by the total number of presences of the species; in the context of this method, 'species' means the abundance class of a given species, e.g. *Acartia clausi* (2) is the second abundance class of the species *Acartia clausi*:

$$FD_{i,k,v} = F_{i,k,v} / FT_i$$

where $FD_{i,k,v}$ = the fidelity of a species abundance class i in class k of variable v ; $F_{i,k,v}$ = the frequency (number of presences) of species abundance class i in class k of variable v ; and FT_i = the total frequency (number of presences) of species abundance class i . Fidelity may be estimated for the classes of either a quantitative or a qualitative descriptor. The ecological or geographic behaviour of a species will be defined by the set of fidelity vectors

related to each ecological or geographic variable. However, since species may be absent from the sites corresponding to some classes of the variable of interest, cumulative histograms of fidelity will be used in the following treatments, cumulated over the ordered classes of the ecological or geographic variables; this solution could not be used with unordered ecological descriptors, of course. This is a way of minimising sampling error when the range of variation of an ecological variable has been incompletely sampled.

Ecological distance among species (D)

The distance among species in ecological space is computed as the weighted Euclidean distance between the vectors of cumulated fidelities (from left to right) associated with all ecological or geographic variables:

$$D_{i,j} = \sqrt{\sum_{m=1}^P \sum_{k=1}^{CL} (GD_{i,k,m} - GD_{j,k,m})^2 / \sum_{m=1}^P CL_m}$$

where $GD_{i,k,m}$ = the cumulated fidelity of species i for class k of variable m ; $GD_{j,k,m}$ = the cumulated fidelity of species j for class k of variable m ; CL = the number of classes for variable m ; and P = the number of variables.

Results differ from those of classical distance functions. With the usual Euclidean distance formula, for

instance, a large distance would be associated with 2 species limited to different and non-overlapping classes in the scale of variation of a variable, while in the above formula, which takes into account the cumulated fidelities, species located near one another along a gradient have a smaller distance than species located farther away along the gradient.

Hydrological position and geographic position

The climatic position (CP) proposed by Brisse & Grandjouan (1978) may be transposed to the concepts of hydrological position (HP; space of the environmental variables) and geographic position (GP; space of the geographic coordinates). It indicates the relative position of a species compared to the distribution of all species along a gradient or transect, divided into CL classes. The HP (in %) of a species *i* for an ecological variable *v* is the covariation index given by:

$$HP(i,v) = 100 \left(\sum_{m=1}^{CL} \sum_{k=1}^{CL} SGN(k-m) F_{i,k,v} \times FK_{m,v} \right) / FT_i \times FX$$

where $SGN(k-m)$ is the sign of $k-m$; $F_{i,k,v}$ = the frequency of species *i* for class *k* of variable *v*; $FK_{m,v}$ = the total number of observations in class *m* of variable *v*; FX = the total number of observations; and FT_i = the total frequency of species *i*. The GP (in %) is defined in the same way.

This index varies between -100 % and +100 %. A null value for HP indicates that the species occupies the median position along the gradient for the whole set of species. If HP reaches 50 %, the species has a position half-way in the upper part of the distribution of all species. HP is more sensitive than the median or the mean, because it takes into account all the pairwise comparisons between species.

Indicative value (IV)

The indicative value [IV(*i,v*), in %] of species *i* measures its relative frequency in the various classes of a variable *v*, compared to that of the whole set of species. It indicates the attraction of a species to a specific part of the variation scale of a variable. It is calculated by the ratio of 2 standard deviations:

$$IV(i,v) = 100 \sqrt{VX / V(i,v)} - 100$$

where VX = the variance of the frequency distribution of all species among the classes of variable *v* weighted by the total frequency of occurrence of all species, and $V(i,v)$ = the variance of the frequency distribution of

species *i* among the classes of variable *v* weighted by the total frequency of occurrence of species *i*. To determine a variance weighted by the frequencies, for instance $V(i,v)$, it is necessary to use the formula:

$$V(i,v) = \left(\sum_j f_{i,j} v_j^2 - \frac{\left(\sum_j f_{i,j} v_j \right)^2}{\sum_j f_{i,j}} \right) / \sum_j f_{i,j}$$

where v_j = the value attributed to the class *j* of variable *v*, and $f_{i,j}$ = the number of individuals of species *i* found in the class *j* of variable *v* (e.g. the weights defined here by the frequencies). For VX , in this last expression, $f_{i,j}$ corresponds to the frequency for all the species.

IV(*i,v*) cannot be negative. If IV(*i,v*) has a null value, the species is not indicative, its distribution along variable *i* being identical to that of all species. A value of 100 % indicates that the species is twice as concentrated (or less dispersed) than the whole set of species along that gradient. If the species is represented by a single class of the variable, $V(i,v)$ is zero and by convention the value is set to 1000 %.

Combining indices IV and HP provides important ecological information (see examples in Fig. 8, below). A species with HP near 0 and low IV is distributed like the whole set of species (which may or may not be significantly correlated with one of the environmental variables). Low HP combined with high IV shows that the species is indicative of a few classes close to the mode of the distribution of all species. Notice, however, that the coverage of the sampling program affects the values of IV. If a species has not been sampled along the full extent of its variation along a given variable, its IV value will be low, whereas if the sampling has a greater extent than that species' scale of variation, IV becomes very high.

To simplify the graphical representation of the ordinations, we arbitrarily decided to retain only the classes of species abundance with IV values higher than 25 % for at least one of the environmental variables. Thus only 38 abundance classes were kept instead of the original 54. This procedure eliminated inadequately sampled species, as well as those that have the same behaviour as the whole set of species.

Distance among observations

Ordinary ecological (or geographic) distances among observations are usually computed from the measured values of the environmental variables, or from the geographic coordinates. By doing so, we determine some groups of stations from the physical (or climatic) frame, even if that classification does not present any eco-

logical meaning. In the phytosociological method, distances among stations take the variables into account, but only if they involve an ecological response of the species. Each station is seen as a virtual species with cumulated fidelity, for each environmental variable, corresponding to the mean of the cumulated fidelity values (mean of the $GD_{i,k,m}$ values) of the species present in that station. Then weighted distances are calculated for all pairs of stations. Since the reference space remains unchanged, it is possible to represent both the species and the stations in the same ordination; distances between species and stations are computed as described here, since stations correspond to virtual species. In the same way, values of HP and IV can be estimated for each station (observation point). They indicate the position of each station along the gradients and point out the most ecologically (or geographically) meaningful stations.

Biological space

Biological fidelity (FB)

The relationship among species is not estimated with ordinary correlations, but rather by using an approach similar to the concept of fidelity for a species. Thus the fidelity of species i with respect to species j is the relative frequency of i in the cases where j is present. Let a be the number of stations containing i and j , b the number of stations containing i only, and c the number of stations containing j only; we then have:

$$FB(i,j) = \frac{a}{a+b} \text{ (fidelity of } i \text{ with respect to } j), \text{ and}$$

$$FB(j,i) = \frac{a}{a+c} \text{ (fidelity of } j \text{ with respect to } i).$$

Thus there are 2 asymmetrical fidelities for 2 species. For a given species, the higher the frequency of common occurrence is compared to its overall distribution, the higher is the fidelity. Thus it is possible to estimate an asymmetrical matrix of fidelities among species. The rows correspond to the fidelities of one species to all the others, the columns to the fidelities of all species to the given one.

Biological distance among species

Contrary to classical methods, the distance among species in biological space is not computed from their observed frequencies; instead it is defined as the Euclidean distance among the fidelities of the 2 species with respect to all the other species, i.e. among the rows of the fidelity matrix (matrix of $F(i,j)$ values com-

puted in the previous subsection). So it is mathematically possible for 2 species to be identical (null distance), even if they have never been found together at the same station! That distance refers to the concept of community, and not to that of co-occurrence in a geographic or environmental space.

Biological distance among observations

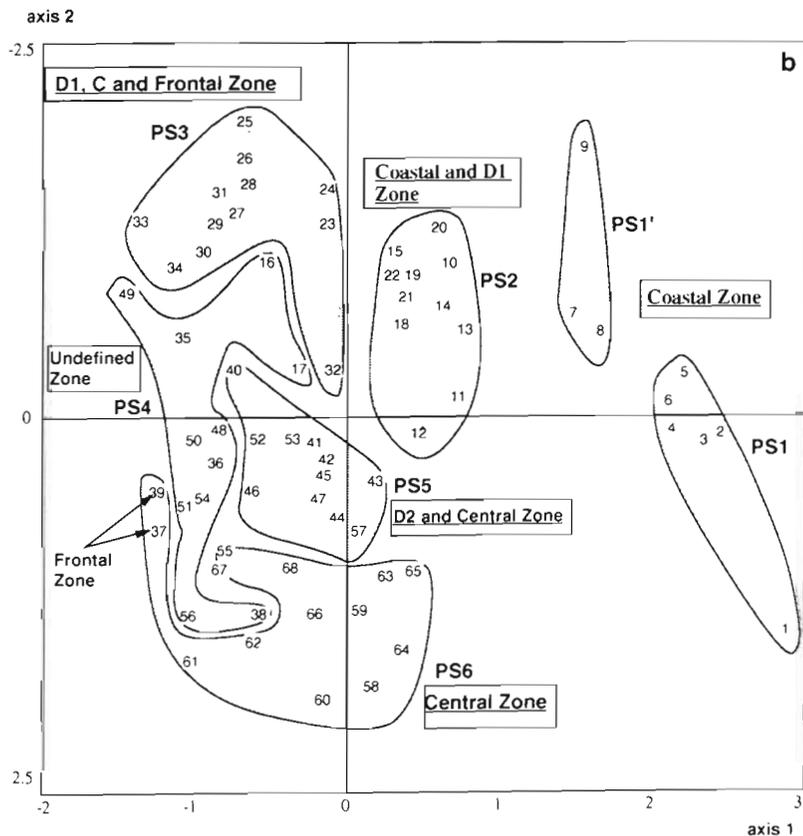
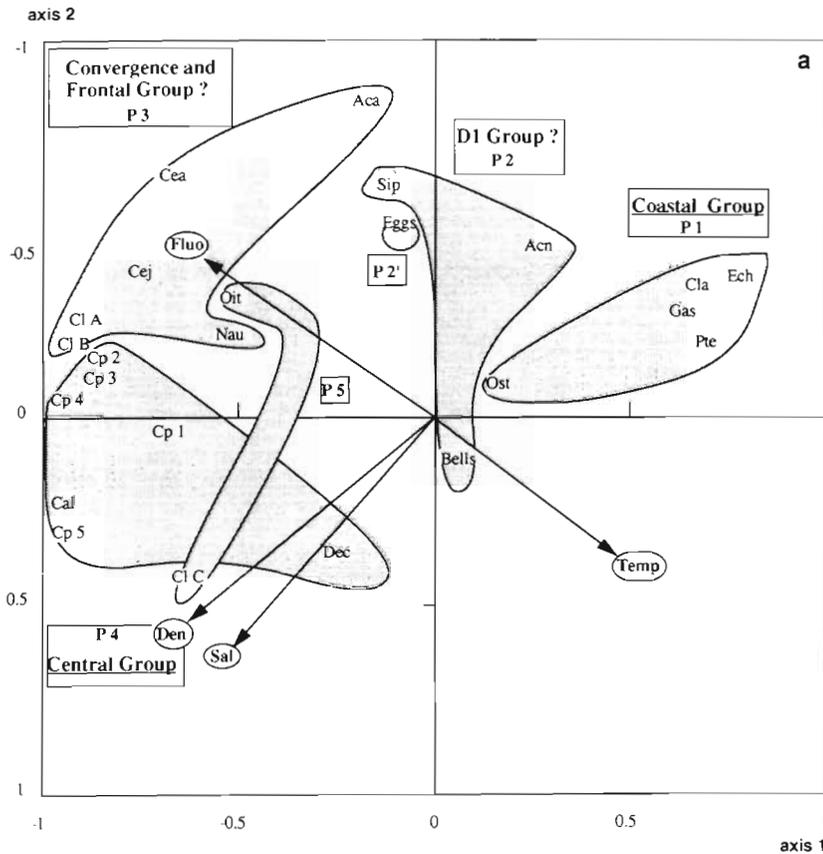
As in the ecological space, the biological distance between 2 stations is defined as the Euclidean distance between 2 virtual species with fidelities corresponding to the vectors of the means of the fidelities of the species present at these stations. As was the case with ecological distances, species and observations can be represented in the same biological space, since the referential space is the same, and since species-observation distances can be computed in the same way as distances among species, as explained above.

RESULTS OF CLASSICAL PROCEDURES

The ordination and classification results obtained by the classical methods (PCA and CA) are shown in Figs. 6 & 7. Here only the first factorial plane (axes 1 and 2) is represented and commented upon. In each ordination, groups have been defined by clustering from a matrix of Euclidean distances among species, or among observations using their coordinates in the factorial space (5 axes).

PCA: the species space

Fig. 6a shows the correlations between the first 2 principal components, on the one hand, and the 24 log-transformed species abundances and the 4 environmental variables (arrows, included as supplementary variables), on the other. The first axis accounts for 30 % of the total variance, the second for 21 % (together: 51 %). Two environmental gradients, temperature opposite to fluorescence, and the salinity-density group, appear uncorrelated. They correspond to directions at 45° of the principal axes. The cloud of the species forms a 'horseshoe'. The species vectors are well represented by that plane, with only 2 poorly represented species (Ost = ostracods and Bells = bells of calyphores, which are close to the centre and have low weights in this 1-2 axis plane), although no species is especially responsive to high temperature, and high salinity corresponds to only 1 species (CIC = *Clausocalanus* spp. C).



The first group of species identified by clustering (P1) includes the most coastal organisms (echinoderm larvae, cladocerans, gastropods, pteropods and ostracods) and seems characteristic of waters with high temperatures, low chlorophyll and low density and salinity. Fig. 6b, which presents the ordination of the stations in the first factorial plane, indicates that the species of the P1 group are found in the first part of the transect: Stns 1 to 9. All these species belong to the peripheral zone, one of the hydrological compartments defined in Fig. 2.

Group 2 (P2) contains 3 taxa: siphonophores, acantharians and calycophore bells (crustacean eggs form an isolated category, P2', located close to P2). This group is situated near the origin of the system of axes, so it is difficult to assign to it any clear relation with the 4 environmental variables. Fig. 6b may indicate that its distribution extends from Stn 10 to Stn 22, and that it is mostly related to the first D1 divergence. In fact, however, the highest observed abundances of this category are in the frontal zone, between Stns 30 and 40 (Fig. 4). This error shows how dangerous it is with PCA to relate the scaling of the observations to the graph of the descriptors.

Group 3 (P3) is represented by *Acartia clausi*, *Centropages typicus*, the copepod nauplii and *Clausocalanus* spp. A and B. Its distribution corresponds to that of fluorescence and is characteristic of low temperatures and high fluorescence. It corresponds mostly to Stns 23 to 34 (group PS3 in Fig. 6b), i.e. the convergence zone and the beginning of the frontal zone. However, Fig. 4 shows that all these taxa are also abundant (or even present maxima) in the central zone.

Fig. 6. (a) PCA: correlations of the 24 taxonomic groups (see Table 1 for abbreviations) and 4 environmental variables with the first 2 axes. (b) PCA scaling of the 68 observations on the first principal plane (axes 1 and 2). Classification results are represented by envelopes. (P and PS are abbreviations denoting the different groups in the PCA)

Group 4 (P4) contains the copepodite stages and adults of *Calanus helgolandicus* and the decapod larvae. They are associated with high salinity and density, and with the last stations of the transect (50 to 68), mixed with erratic observations from the frontal zone (35, 37, 39). Here, again, the description of the environmental frame is not very good, because these stations are not at all homogeneous in terms of their hydrological conditions (Fig. 3).

The classification algorithm put *Oithona* sp. and *Clausocalanus* spp. C together in group P5, which seems very dispersed in the ordination. No preferential stations can be assigned to it. Fig. 4 shows the distribution of these species to be mostly centred in the frontal zone, a phenomenon which is not made clear by the PCA ordination.

At this step one may conclude that PCA produces an unclear ecological representation of the species succession along the gradient under study. The species groups are only partly obvious. Relationships with the environmental variables are often fuzzy, and the grouping of stations, which does not preserve the temporal contiguity along the transect, seems difficult to relate to the distribution of the species groups. Neither the correlation space nor the scaling space is able to lead to a comprehensive interpretation. Taking into account the next principal axes of lower variance (3rd, 4th, etc., not presented here) does not help clarify the situation – quite the contrary.

CA: the dual observation-descriptor space

CA was performed on a binary contingency table containing 68 observation rows and 79 descriptor classes. The thresholds for the abundance class descriptors are the same as in the phytosociological method, except that here 2 thresholds define 3 classes, as explained in the 'Methods', whereas only 2 binary variables are constructed in the phytosociological method. For the environmental variables, 3 classes were used. Although species, environmental factors and observations may be represented in the same graph in the CA method, 2 graphs (Fig. 7a, b) were drawn in order to reduce confusion.

In Fig. 7a, group F1 contains the high-abundance classes for the coastal species (echinoderm larvae, cladocerans, gastropods, pteropods), with the lowest salinity and density classes. This condition corresponds to the coastal waters; Fig. 7b confirms that this group is located in the more coastal stations (1 to 13).

Group F2 (Fig. 7a) contains the lowest levels of almost all the 24 taxonomic groups, associated to the highest classes of temperature and the first 2 classes of fluorescence. Fig. 7b indicates that this heterogeneous

group is found in the peripheral zone (Stns 14 to 23) as well as near divergence zone D2 (Stns 42 to 47) and in the central zone (Stns 63 and 64). This result is confirmed in Fig. 4.

Groups F3 and F4 contain the high classes of *Calanus helgolandicus* copepodites 2, 3 and 4, crustacean eggs, *Acartia clausi*, *Centropages typicus*, and *Clausocalanus* spp. A and B. The proximity of the high class of fluorescence (Flu3) and of the lowest temperature class (Tem1), and the correspondence with Stns 24 to 34 in Fig. 7b, show that these taxonomic groups are characteristic of the frontal zone. Surprising is the position of high classes of *A. clausi*, which have rather low abundance in the frontal zone (Fig. 4).

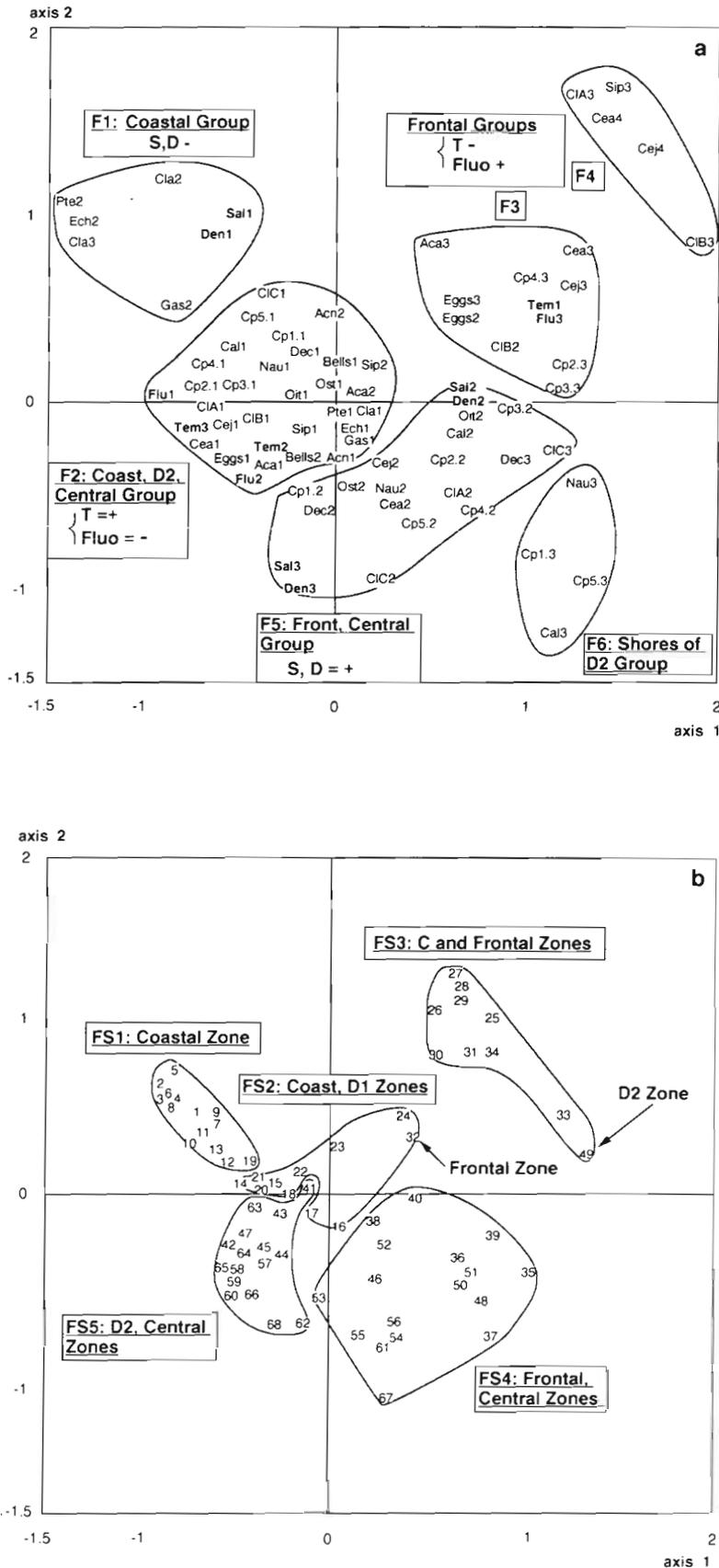
Group F5 includes the intermediate levels of several taxonomic groups and the highest levels of the decapod larvae and of *Clausocalanus* spp. C. Taking into account the presence of high salinity and density (medium and high classes), as well as the corresponding stations (Fig. 7b), this group shows species which extend from the frontal to the central zone, but are absent or rare in divergence zone D2; that distribution is confirmed in Fig. 4.

The last group, F6, contains only level 3 of *Calanus helgolandicus* copepodites 1 and 5, *C. helgolandicus* adults and copepod nauplii. These taxonomic groups are found near the limits of divergence zone D2 (Stns 35 to 39 and 48, 50 and 51, according to Fig. 7b). Considering Fig. 4, however, the highest levels of *C. helgolandicus* adults and copepodites 5 are clearly not found at the same stations.

Thus correspondence analysis separates coastal taxonomic groups from frontal groups (gradient represented by the first axis), while axis 2 opposes both coastal and frontal taxa to those of the central zone. The results are much more detailed than with PCA, since they specify the distribution of the abundance classes as well as their position in the range of variation of the environmental variables. Some species assemblages seem rather spurious, however; only the classes characterising delimited hydrological compartments are clearly interpretable. The classification method used for the stations, which does not preserve their spatial contiguity, is one reason for this ambiguity.

RESULTS OF THE PHYTOSOCIOLOGICAL METHOD

Because it separates the information into 3 distinct spaces, the phytosociological method yields more precise information about each descriptor and each observation. The abundance classes of the taxonomic groups are noted below in parentheses.



Geographic space

Fig. 8 shows the fidelities of all the taxonomic groups combined across 11 geographic classes. Fidelities are here computed as the mean of the fidelities of all species in each geographic class. The mode of maximum density is located near the middle of the transect, between Stns 26 and 38 (abundance classes with null geographic position index, GP, and corresponding null indicative value, IV). The taxonomic groups essentially concentrated in the frontal zone have a GP index close to zero [examples in Fig. 8 are *Centropages typicus* (3) and *Clausocalanus* spp. B (2)], whereas the coastal and central taxonomic groups present higher GP values [*Clausocalanus* spp. C (1), echinoderm larvae (2); Fig. 8]. Table 2 shows the taxonomic abundance classes with the highest indicative values (IV higher than 150 %) along the geographic transect, as well as their GP indices. The high abundance classes of the coastal taxa (cladocerans, echinoderm larvae, pteropods) have IV values near 500 %, which points out a very strong dependence on their spatial distributions. Taxonomic groups of the front and of the convergence zones [*Clausocalanus* spp. A (2), *Centropages typicus* (2) and (3), siphonophores (3)] also depend upon a geographic region (IV values: 200 to 300 %), located in the middle of the transect. Notice the spatial position of *Clausocalanus* spp. B (2) in Fig. 8, which is shifted offshore compared to the other frontal species, with low concentration in the convergence zone area ('C' in Fig. 3).

Table 3 gives the GP and IV index values for 14 stations selected in the ecological space (see following section).

Fig. 7. (a) CA: scaling of the 67 abundance classes (see text: 'Numerical analyses') and of the 12 classes of the environmental variables in the first inertial plane (axes 1 and 2). (b) CA: scaling of the 68 observations in the first inertial plane. Classification results are represented by envelopes. (F and FS are abbreviations denoting the different groups in the CA)

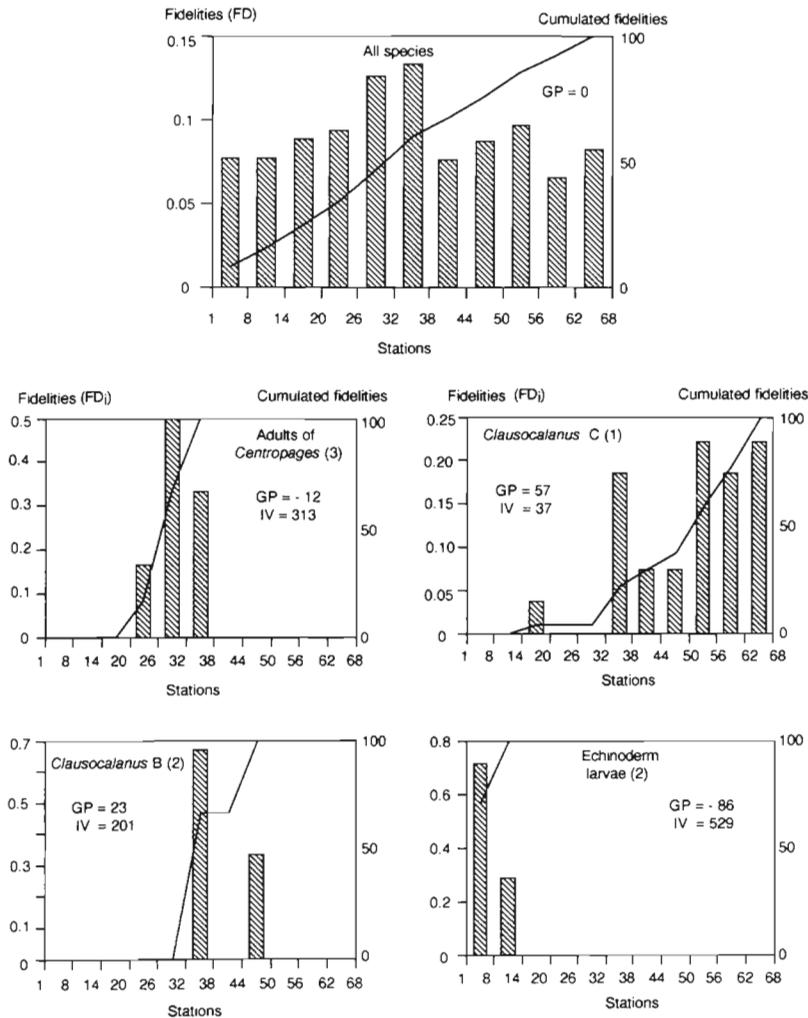


Fig. 8. Geographic space: fidelities (FD; histogram), cumulated fidelities (line) and geographical positions (GP) for all taxonomic groups, as well as for 4 chosen taxonomic abundance classes (with their indicative values, IV)

Table 2. Geographic position (GP) index of the most indicative taxonomic abundance classes (those with IV > 150 %) in the geographical space

| Taxonomic group | Geographic variable | |
|---------------------------------------|---------------------|-----|
| | IV | GP |
| Cladocerans (3) | 480 | -81 |
| Echinoderms (2) | 529 | -86 |
| Pteropods (2) | 480 | -84 |
| Gastropods (2) | 176 | -69 |
| <i>Clausocalanus</i> spp. A (2) | 160 | -13 |
| <i>Centropages typicus</i> adults (2) | 170 | -5 |
| <i>C. typicus</i> adults (3) | 313 | -12 |
| <i>C. typicus</i> juveniles (3) | 201 | -7 |
| Siphonophores (3) | 290 | -5 |
| <i>Clausocalanus</i> spp. B (2) | 201 | +23 |

The observations with highest IV indices for the geographic variable (26, 32, 41, 61, 64) belong to the frontal or central zones. High abundances are found at these stations for several taxonomic groups: *Acartia clausi* and *Clausocalanus* spp. A at Stn 26; *Oithona* sp. and *Clausocalanus* spp. B and C at Stn 32; crustacean eggs at Stn 41; *Clausocalanus* spp. C and *Calanus helgolandicus* copepodites 4 and 5 at Stn 61; and acantharians at Stn 64.

Fig. 9 shows the classification (envelopes) and the ordination by MDS, based upon the distances computed among the 38 selected abundance classes (see section on the phytosociological method) and the 14 stations listed in Table 3. The descriptors with high IV values in this geographic space are represented in special fonts (see figure legend). Axis 1 shows mainly the succession of the populations along the transect, whereas axis 2 untangles the heterogeneity of the frontal zone regarding the ordination of the species.

G1 is a group of species presenting their maxima in the coastal zone, as noted above. G2 contains intermediate taxonomic groups (low level for gastropods and the 2 levels of *Acartia clausi*) located between the peripheral and frontal zones, i.e. on the first convergence with GP values between 25 and -20. All the taxonomic groups of G3 present high abundances in the frontal zone. Species below the first axis [*Centropages typicus* adults (2) and (3), *Clausocalanus* spp. A (2)] are abundant in the frontal as well as in the convergence zone. Species above the first axis [*C. typicus* juveniles (3), *Oithona* sp., ostracods and siphonophores (3)] are concentrated mostly in the frontal zone, while they remain low or may even be absent in the convergence zone. This differentiation among the frontal species, displayed by axis 2, leads to the distinction of G5 [*Clausocalanus* spp. B (2), copepod nauplii (2) and crustacean eggs (2)], which presents high densities in both parts of divergence D2 [justified by the indicative GP of *Clausocalanus* spp. B (2) and of Stn 41]. Notice that Stn 41 corresponds to relatively poor densities, except for the crustacean eggs (18 % of the organisms are eggs at that station, although on average they represent only 1.5 % of the total abundances in all sta-

Table 3. Geographical position (GP) indices and indicative values (IV) of the 14 stations selected in the ecological space

| Station | Geographic variable | |
|---------|---------------------|-----|
| | IV | GP |
| 2 | 18 | -59 |
| 3 | 18 | -59 |
| 4 | 16 | -57 |
| 6 | 14 | -58 |
| 12 | 8 | -51 |
| 26 | 34 | 3.6 |
| 32 | 43 | 8 |
| 34 | 27 | 17 |
| 36 | 29 | 25 |
| 41 | 44 | 9 |
| 46 | 29 | 22 |
| 61 | 31 | 34 |
| 63 | 0 | 20 |
| 64 | 38 | 55 |

tions). Group G4 contains several low-abundance classes of *Calanus helgolandicus* copepodites and *Centropages typicus* juveniles. Stns 26 and 32 are indicative of the frontal zone. Group G6 contains taxonomic groups that are abundant in the frontal as well as in the entire central zone (indicative stations: 61 and 64), but rare in the D2 divergence. This lack of spatial contiguity explains why the taxonomic groups were never found to be indicative. Group G7, below axis 1, contains Stn 63 and level 2 of the calyphore bells. This taxonomic group is abundant in the frontal and central zones, but also in the convergence area and the off-shore stations of the peripheral zone.

So we have seen that the analysis in geographic space was able to detect the main population changes along the transect; of course, the method would be much more powerful if the data were positioned with reference to a multidimensional space of coordinates (depth, hours, date, latitude, etc.), as is done in the following section. A more precise description would require us to consider the fidelity diagrams for all the taxonomic groups presenting high IV values, and to tabulate the correspondence between the GP index values and the samples' geographic positions.

Ecological space

Considering all species abundance classes, the variability is not identical for all 4 environmental variables

(Fig. 10, top row). For temperature and fluorescence (Fig. 10a), fidelities show trends which are generally decreasing or increasing, opposite to one another, which is a consequence of the negative correlation found between these 2 variables. Thus the optimal conditions of temperature and fluorescence, most favourable to all the organisms, have probably not been found. Colder water, with higher primary productivity, could lead to higher zooplankton biomass. The highest fidelities correspond here to classes of temperature lower than 13 °C, and classes of fluorescence higher than 1.4 mg C m⁻³. In contrast, the response to salinity and density (Fig. 10b) exhibits a unimodal shape. The mode is between 38.20 and 38.35 PSU in salinity, and from 28.89 to 29.10 (kg m⁻³ - 1000) in density. Some taxonomic abundance classes with HP near zero present the same behaviour as all the abundance classes together: *Acartia clausi* (1) for temperature and fluorescence (Fig. 10a, second row), and *Centropages typicus* juveniles (3) for salinity and density (Fig. 10b, second row). Abundance classes with distributions that differ from the whole set of species present higher HP values (as absolute values), e.g. echinoderm larvae for salinity and density. Table 4 contains the abundance classes having at least 2 IV values higher than 200 %. *Clausocalanus* spp. A (2) and B (2), *C. typicus* adults (2) and (3), the siphonophores (3) and *C. typicus* juveniles (3) present high

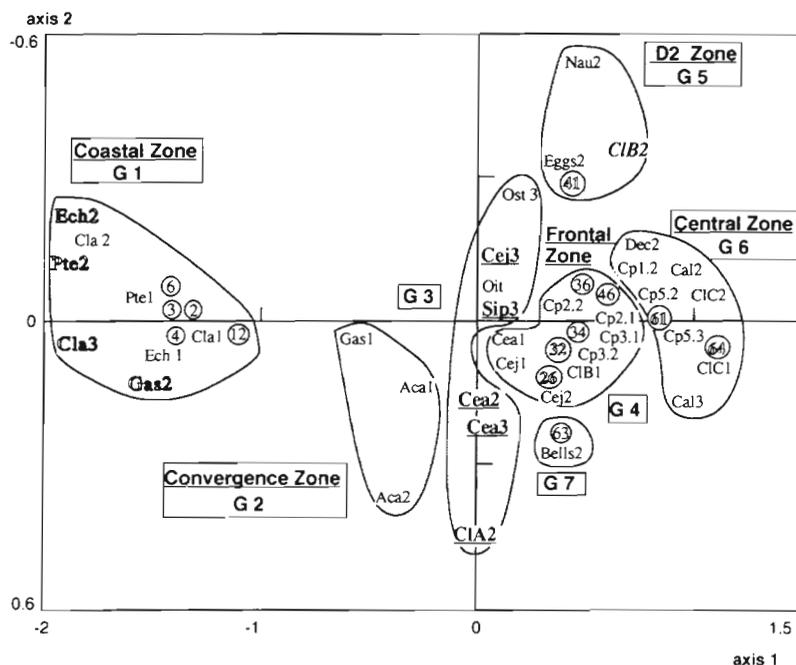


Fig. 9. Geographic space: MDS axes 1 and 2. Stations are represented by circles. Special typestyles (bold, italic, etc.) reflect the most indicative stations or abundance classes; the typestyles differ according to the GP (see Tables 2 & 3). Classification results are represented by envelopes

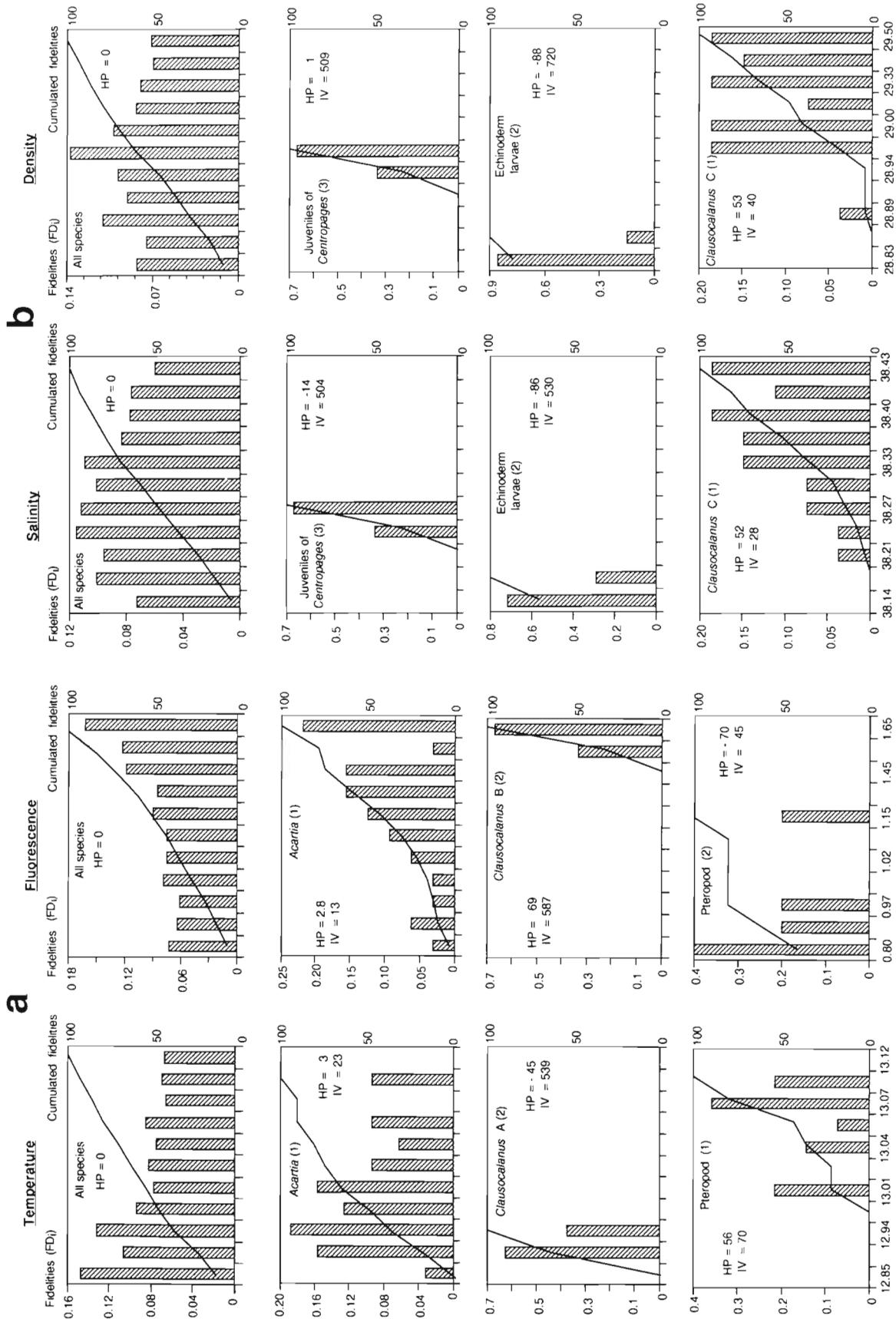


Fig. 10. Ecological space: fidelities, cumulated fidelities, and hydrological position (HP) with respect to (a) temperature (°C) and fluorescence (mg chl a l⁻¹) and (b) salinity (PSU) and density (kg m⁻³ - 1000), for all taxonomic groups together, and for some characteristic abundance classes [with their indicative values (IV) shown]

Table 4. Hydrological position (HP) indices of the most indicative taxonomic abundance classes (those with IV > 200 %) in the ecological space

| Taxonomic group | Ecological variables | | | | | | | |
|---------------------------------------|----------------------|-----|--------------|-----|----------|-------|---------|-----|
| | Temperature | | Fluorescence | | Salinity | | Density | |
| | IV | HP | IV | HP | IV | HP | IV | HP |
| <i>Clausocalanus</i> spp. A (2) | 539 | -45 | 274 | 66 | 85 | -32 | 82 | -26 |
| <i>Clausocalanus</i> spp. B (2) | 228 | -66 | 587 | 69 | 128 | -0.06 | 204 | 24 |
| <i>Centropages typicus</i> adults (2) | 238 | -52 | 227 | 55 | 90 | -19 | 93 | -10 |
| <i>C. typicus</i> adults (3) | 279 | -56 | 261 | 55 | 314 | -26 | 169 | -16 |
| Siphonophores (3) | 324 | -68 | 363 | 58 | 128 | -32 | 130 | -15 |
| <i>C. typicus</i> juveniles (3) | 228 | -65 | 297 | 50 | 504 | -14 | 509 | 1 |
| Ostracods (3) | 28 | -42 | 39 | 21 | 350 | 15 | 353 | 9 |
| Cladocerans (2) | 34 | 44 | 7 | -57 | 472 | -82 | 337 | -82 |
| Cladocerans (3) | 68 | 45 | 54 | -68 | 481 | -79 | 487 | -83 |
| Echinoderms (2) | 85 | 46 | 45 | -66 | 530 | -86 | 720 | -88 |
| Pteropods (2) | 73 | 42 | 45 | -70 | 612 | -87 | 617 | -87 |

concentrations correlated with high values of fluorescence and low values of temperature (frontal characteristics). Differences exist between these species, however: *Clausocalanus* spp. A (2) mostly depends on the low temperatures, *Clausocalanus* spp. B (2) on high fluorescence, and *C. typicus* juveniles (3) on mean values of salinity and density. Ostracods are indicative of intermediate classes of salinity and density. Cladocerans (2) and (3), echinoderm larvae (3) and the pteropods (2) are mostly concentrated in the lowest classes of salinity and density (coastal zone). Notice that only the upper levels of abundance (2 or 3) of these taxonomic groups have large indicative values (Table 4).

The HP and IV values for the stations are shown in Table 5 (IV > 23 % for at least one of the ecological variables). Stns 2, 3, 4, 6 and 12 indicate high temperatures (peripheral water), but are distributed in classes of low fluorescence, salinity and density (Fig. 3). Stns 26, 32 and 34 display high fluorescence

(frontal water) and intermediate values for density. Stns 36 and 46 present higher IV values for salinity and density. The high IV values for Stns 41 and 63 only indicate low temperature and low fluorescence, respectively. Stns 61 and 64 are characterised by high salinity and density (central zone).

Fig. 11 shows the ordination and the classification results for the 38 species (see section on the phytosociological method) and the stations retained.

Group E1 contains Stns 2, 3, 4, 6 and 12, which are all indicative of high temperatures, with taxonomic groups such as cladocerans, echinoderm larvae, pteropods and gastropods. The higher abundance classes of these taxonomic groups are indicative of low salinity, density and fluorescence, and of high temperature. E1 corresponds to the coastal biotope.

Group E2 contains the first abundance classes for adults and juveniles of *Centropages typicus*, *Clausocalanus* spp. B, *Calanus helgolandicus* copepodites,

Table 5. Hydrological position (HP) index of the most indicative stations (those with IV > 23 %) in the ecological space

| Station | Ecological variables | | | | | | | |
|---------|----------------------|-----|--------------|-----|----------|-----|---------|-----|
| | Temperature | | Fluorescence | | Salinity | | Density | |
| | IV | HP | IV | HP | IV | HP | IV | HP |
| 2 | 30 | 41 | 10 | -48 | 7 | -48 | 8 | -53 |
| 3 | 30 | 41 | 10 | -48 | 7 | -48 | 8 | -53 |
| 4 | 28 | 41 | 8.5 | -47 | 6 | -47 | 6 | -51 |
| 6 | 28 | 40 | 7.5 | -47 | 4 | -47 | 0 | -52 |
| 12 | 24 | 40 | 7 | -44 | 0 | -44 | 4 | -44 |
| 26 | 18 | -22 | 31 | 27 | 20 | -2 | 27 | 0.2 |
| 32 | 11 | -29 | 37 | 30 | 23 | 0.1 | 34 | 4 |
| 34 | 3 | -21 | 20 | 23 | 16 | 9 | 25 | 13 |
| 36 | 5 | -16 | 12 | 14 | 17 | 21 | 24 | 23 |
| 41 | 36 | -34 | 0 | 17 | 10 | 0.5 | 6 | 9 |
| 46 | 5 | -13 | 14 | 15 | 15 | 19 | 23 | 20 |
| 61 | 4 | -14 | 8 | 9 | 20 | 30 | 27 | 32 |
| 63 | 0 | 23 | 25 | -25 | 0 | 20 | 0 | 18 |
| 64 | 0 | 10 | 10 | -8 | 30 | 47 | 41 | 49 |

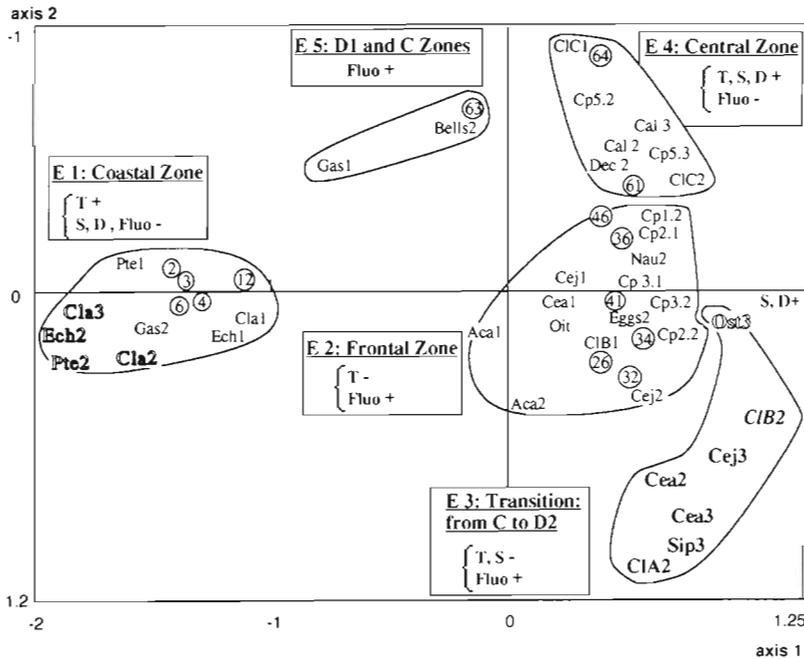


Fig. 11. Ecological space: MDS axes 1 and 2. Stations are represented by circles. Special typetypes (bold, italic, etc.) reflect the most indicative stations or abundance classes; the typetypes differ according to the HP (see Tables 4 & 5). Classification results are represented by envelopes

Acartia clausi and *Oithona* sp., and the second level for *A. clausi*, *C. typicus* juveniles, *C. helgolandicus* copepodites 2 and 3, nauplii and crustacean eggs. Their IV values (Table 4) are not especially remarkable, but the presence in the same group of Stns 26, 32, 34, 36, 41 and 46 indicates that E2 corresponds to a frontal group with high fluorescence and low temperature.

Group E3 represents the lowest abundance classes of *Centropages typicus* adults and juveniles, *Clausocalanus* spp. A and B, ostracods and siphonophores. Except for the ostracods, these are all indicative of high fluorescence and low temperature conditions. Their HP values for salinity and density show that these abundance classes are not distributed identically in the frontal zone: the *C. typicus* adults (3) and *Clausocalanus* spp. A (2) present negative HP values, and correspond to salinities between 38.2 and 38.3 PSU and between 28.9 and 28.97; on the other hand, *Clausocalanus* spp. B (2), with higher HP values, is concentrated between 38.3 and 38.33 PSU in salinity, and between 29 and 29.16 in density. E3 corresponds to a transition zone towards the frontal zone (from the convergence to divergence zone D2).

Group E4 contains *Clausocalanus* spp. C and the high-abundance classes of *Calanus helgolandicus* adults and copepodites 5, and of the decapod larvae. IV values are never found to be higher than 150 % but the proximity of Stns 61 and 64 indicates that E4 is

characteristic of the central zone: higher temperature than in the frontal zone, high salinity and density, and low fluorescence in the most offshore part of the transect.

Group E5 contains gastropods (1), calyophore bells (2), and Stn 63. Contrary to the higher gastropod classes, which are confined to the coastal zone, the first class (denoting presence of the organism, in any abundance) is more ubiquitous. It vanishes only above the D1 divergence and near the frontal convergence. The same behaviour is observed for calyophore bells [notice that siphonophores (3) is indicative of the frontal zone E3]. This group, E5, is intermediate between the peripheral (E1) and central zones (E4). So even if the first abundance classes are less indicative of hydrological discontinuities, they can sometimes be informative with regard to the overall distribution of the species.

Biological space

The relationships between species generate more complex ordinations and classifications than those related to the geographic or ecological spaces, because they are the result of 3 correlated processes: spatio-temporal, ecological, and biological-ethological. Thus in Fig. 12, which shows the ordination of the 38 abundance classes and the 14 stations (retained in the ecological space), it may be easy to recognise the coastal group in S1, but the affinities among species of the frontal and central zones are more difficult to see. The set of frontal species is divided up into groups S2, S3, S4 and S5. Some associations are strong: *Acartia clausi* (1) and (2) with *Clausocalanus* spp. A (2) in group S2, *Clausocalanus* spp. B (2) with *Centropages typicus* adults and juveniles (3) and the siphonophores (3) in group S4 (an association also found in the ecological space), and nauplii (2) with crustacean eggs (2) and Stn 41 in group S5 (an association already found in the geographic space).

The central group (E4 in the ecological space) is also split up. *Calanus helgolandicus* adults (2) is close to the frontal stations 36 and 46, *C. helgolandicus* copepodites 5 (2) is close to *Oithona* sp. (which corresponds to a frontal taxonomic group in E2), all of which are mixed in group S6. *C. helgolandicus* adults (3) and *C. helgolandicus* copepodites 1 (2) and 5 (3) become associated in

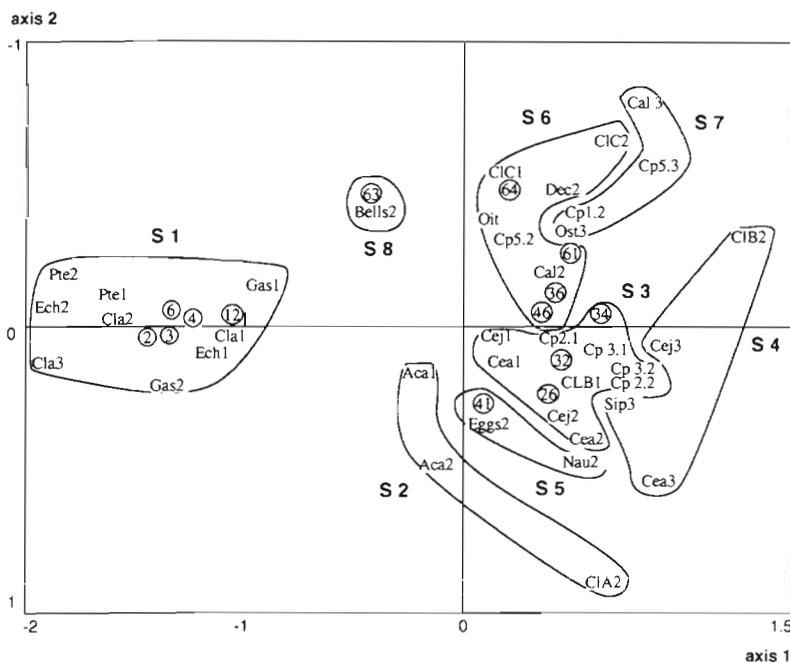


Fig. 12. Biological space: MDS axes 1 and 2. Stations are represented by circles. Classification results are represented by envelopes

group S7 to the ostracods (3) (a frontal category of E3). The calyophore bells (2) are isolated and form group S8 with Stn 63 (found in the ecological space), and remain intermediate between the coastal and offshore zones. The gastropods (1) have been joined to the coastal group in this representation, close to gastropods (2).

The statistical communities detected in the biological space cannot be explained only by the measured environmental variables, nor by the spatial co-occurrence of the species. To explain the associations found in this part of the study, one may have to consider the information provided by the laboratory-based auto-ecological studies found in the literature concerning the trophic and reproductive behaviour of these taxonomic groups, as well as the ability of these organisms to migrate in the vertical and horizontal planes.

Comparison of the three spaces

Mantel tests among the 3 distances matrices among taxonomic abundance classes (1: geographic matrix; 2: ecological matrix; 3: biological matrix) give the following correlations:

$$R_{1,2} = 0.937^{**}; R_{1,3} = 0.858^{**}; R_{2,3} = 0.930^{**}$$

The tests of significance of the Mantel statistics were carried out both by the permutational method and by Mantel's normal approximation. In all 3 cases, the null hypothesis (H_0) of independence between matrices is

rejected at the 1 % significance level (indicated by **). It is more informative, however, to estimate the partial correlations between pairs of matrices and to test for their statistical significance. The partial correlation coefficients and their significance levels are:

$$\begin{aligned} R_{1,2,3} &= 0.736^{**}; \\ R_{1,3,2} &= -0.105^{ns} \text{ (} p = 0.096\text{)}; \\ R_{2,3,1} &= 0.700^{**} \end{aligned}$$

(ns: non-significant Mantel correlation). The $R_{1,3,2}$ value shows that the correlation between geographic and biological distances is caused by their being both correlated with the ecological distances; after removing this effect, there is no relation left between them.

Legendre & Troussellier (1988) have proposed a method to interpret these Mantel and partial Mantel test results in a causal framework. Applying their method, we notice that the following relations are met: $R_{1,2} \neq 0$, $R_{2,3} \neq 0$, $R_{1,2} \geq R_{1,3}$, $R_{2,3} \geq R_{1,3}$, $R_{1,2,3} \neq 0$, $R_{2,3,1} \neq 0$,

$R_{1,3,2} = 0$ (this is the crucial point; it actually means that the partial statistic is not significantly different from zero), $R_{1,2,3} \leq R_{1,2}$, $R_{2,3,1} \leq R_{2,3}$, and finally $R_{1,3} = R_{1,2} \times R_{2,3}$ since $0.937 \times 0.930 \approx 0.858$. The only causal model compatible with these observations is: Geography \Rightarrow Ecology \Rightarrow Biology.

The residual matrix of biological distances after removing the effect of the ecological distances allows another ordination to be created which may be compared to the original one (Fig. 13a, b). The proximity of the taxonomic abundance classes should now depend mostly on their biology and ethology, if one accepts the weak hypothesis that a linear model has managed to completely remove the influence of the ecological factors (Legendre & Troussellier 1988); if this is not the case, then the resulting ordination (Fig. 13b) should neither greatly differ from, nor be clearer than, the original one (Fig. 13a).

Coastal species remain together in the centre of the plane. Axis 1 tends to separate the higher abundance classes of *Calanus helgolandicus* adults and copepodites 1 and 5, and ostracods, from those of *Centropages typicus* adults and juveniles, and *Clausocalanus* spp. B and C. Crustacean eggs and copepod nauplii are also separated, eggs being close to frontal taxa, while nauplii are close to central taxa. Axis 2 recognises 2 types of frontal categories and opposes *Clausocalanus* spp. A (2) and B (2). Siphonophores (3), generally grouped with frontal species, seem more associated with *C. helgolandicus* (2) and *Clausocalanus* spp. B and C.

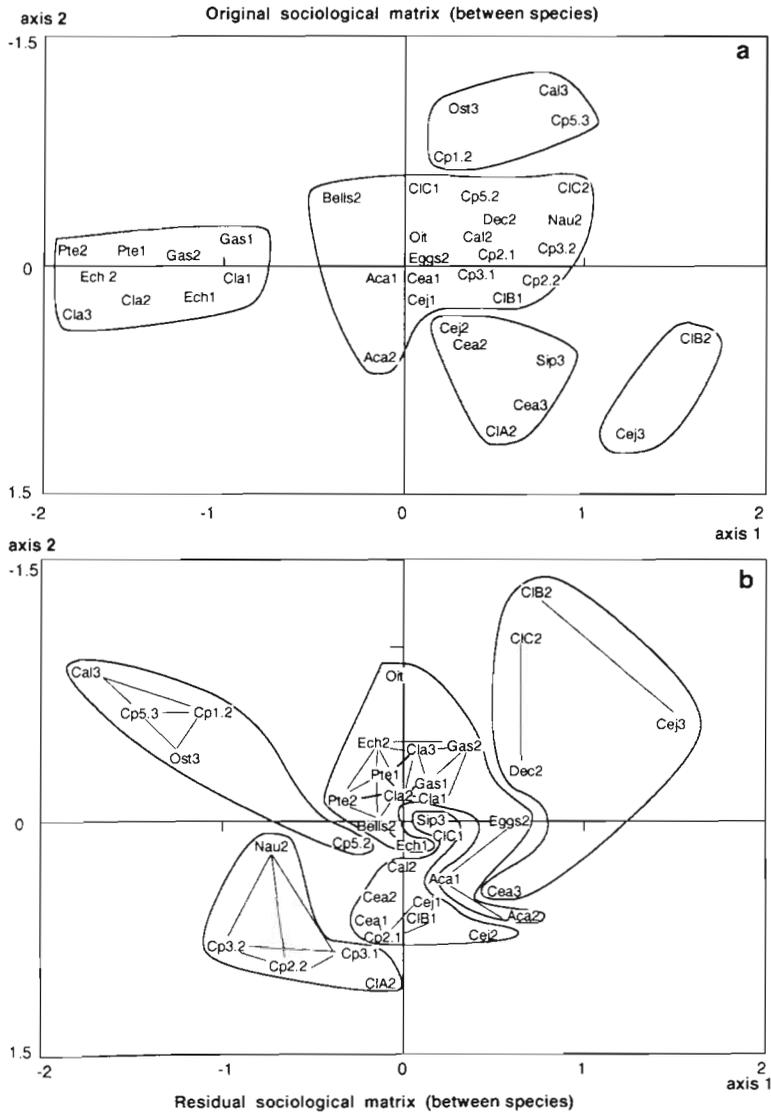


Fig. 13. (a) Original biological space: MDS axes 1 and 2. The ordination was recomputed for the abundance classes only, in order to make it comparable to Fig. 13b. (b) Residual biological space: MDS axes 1 and 2, computed using the residual biological distances, controlling for the linear effect of the ecological distances. The lines joining some taxonomic groups indicate that these abundance classes were already in the same group, in Fig. 13a, before the ecological influence was removed. In both panels, classification results are represented by envelopes

calanus spp. C (1). A more precise interpretation of the ordination could be reached using the following principles:

(1) If species that are close together in the biological space remain associated in the residual space, this means that they present a similar behaviour and similar responses with respect to the ecological variables, e.g. *Calanus helgolandicus* adults (3) and copepodites 5 (3). Taxonomic groups that behave in this way are connected by lines in Fig. 13b.

(2) Abundance classes that become more distant in the residual space were only held together by their sharing the same ecological conditions; this is the case with *Clausocalanus* spp. A and B (2), for instance, which have different sizes, so that they should have different migratory capacities and different trophic behaviours. Siphonophores and the adults of *Centropages typicus* (3) have different trophic needs: the former are carnivorous whereas the latter are herbivorous (and sometimes omnivorous). Crustacean eggs and copepod nauplii do not have the same swimming ability, which implies a different impact of turbulent diffusion.

(3) Taxonomic groups that are closer in the residual space have similar behaviour, but their reactions to the ecological factors are different. *Clausocalanus* spp. B and C (2), for example, are not sufficiently different in size to have distinct ethology. On the other hand, the coastal species, which are represented by the group in the centre of the ordination graph comprising echinoderm larvae, gastropods, pteropods, cladocerans and calycophore bells, have the same ecology and behaviour, but they seem to be isolated from the other taxa mostly by their ecological adaptations.

DISCUSSION AND CONCLUSION

As stated by Gordon (1987), 'human ingenuity is quite capable of providing a *post hoc* justification of dubious classifications'. The same can certainly be said of ordinations, as convincingly shown by the vast ecological literature discussing scaling results. The purpose of this study is not to reject the prevailing multivariate data treatments. Certainly, techniques more sophisticated than simple PCA or CA would provide more interesting information about the frontal ecosystem. These 2 examples permit, however, a better understanding of the fundamental innovations brought about by the phytosociological method proposed by Brisse & Grandjouan.

Partitioning species abundance distributions into abundance classes results in giving each class, and each species, the same weight, and also in establishing indicative levels of significant responses along environmental gradients. The coding performed in classical CA, where the data are coded into a 0-1

(disjunctive) table, differs from that of the phytosociological method in that the classes in the CA coding are the successive abundances; in the phytosociological method, on the other hand, each class is retrocumulative and corresponds to a threshold, a station being coded 1 for that class if the species is more abundant than the threshold. Thus the contrast between the different levels is increased. The key characteristic of the phytosociological method is, however, that instead of defining distances or correlations from species frequencies or abundances, these are computed from cumulated fidelities with respect to external variables. This allows the investigator to take due account of the concentration of an abundance class into one or a few classes of an external variable. PCA and CA did not highlight any particular behaviour of the siphonophores, for instance, whereas the phytosociological method clearly pointed out that this taxon, at level 3, strongly indicates the frontal biotope. The high-abundance class of the ostracods presented a distribution which was much clearer than that observed when considering all frequencies. Differing from the other coastal species which do not cross the D1 divergence (Ibanez 1982, Boucher et al. 1987), gastropod larvae go beyond that divergence when considering abundance class 2 only, while they are perceived as almost uniformly distributed along the transect when taking into account only abundance class 1 (presence-absence).

The HP or GP and the IV index are valuable for interpreting the ordinations. They allow detection of the sensitivity of some abundance classes to some classes of the external variables. Although *Clausocalanus* spp. A and B are not clearly separated in the PCA and CA ordinations, they were found to be distinct in the ecological space because of their relationship with low temperature and high fluorescence, and also because of their precise location in the geographic and biological spaces. While the juveniles of *Centropages typicus*, with high IV values with respect to salinity and density and to geographic position, correspond to a frontal taxonomic group, the distribution of the adults is also important in the convergence zone. HP, GP and IV are functions of the distributions of all the species under study. Of course, the fewer the species, the more difficult it is to interpret the structure of the ecosystem. This notion is not new to ecology: in the same way, diversity indices and rank/frequency diagrams are really meaningful only when all the species have been captured. Considering the stations as 'virtual species' enables them to be represented in the same space as the species and allows their position indices to be estimated. Then it is possible to interpret some non-obvious proximities between stations in the ordination (e.g. Stns 26 and 32 in the ecological space, characterised by high fluorescence and density). It also allows

recognition of stations with some special species assemblage, and serves to detect the locations of the main discontinuities of the environment and/or the ecosystem: in geographic space, Stn 26, selected in Table 3 for its high IV value, corresponds to the coastal limit of the convergence, Stn 32 to its offshore limit, Stn 41 to the offshore limit of the frontal zone, and Stns 61 and 64 are typical of the central zone.

Contrary to the classical methods that offer a representation in a single ordination space, the phytosociological approach separates different ecological matters into distinct ordination diagrams. The geographic space corresponds to the large spatio-temporal gradients. Oceanographic data are characterised by spatio-temporally contagious (i.e. autocorrelated) processes: processes that are active in one location are also active at other points nearby (space), and they have a duration (time). In contrast to ecological space, geographic space takes into account the spatial proximity of the observations. To understand what is going on, the ecologist has to keep in mind the precise distribution of the species: temporal (or geographic) succession, time (or place) of particular discontinuities. The ordination spaces produced by PCA and CA from a species-observation table do not respect the spatio-temporal contiguity of the observations. Other methods have been proposed to perform the segmentation of multivariate ecological data sets: constrained clustering in time (Legendre et al. 1985) or in space (Legendre & Legendre 1984b, Legendre 1987), or constrained scaling (Ibanez 1984). But these methods only aim at partitioning the set of observations into contiguous segments, while the phytosociological method allows groups of taxa that share the same spatio-temporal position to be identified. For instance, *Acartia clausi* and gastropods (1) are detected together precisely in the D1 divergence and in the convergence zone. As noted above, characteristic stations ('virtual species') may be recognised and projected in the same ordination space as the taxa.

The ecological space is founded in the niche concept proposed by Hutchinson (1957): a hypervolume whose coordinates are the values of the external variables. A PCA computed from the environmental variables, followed by projecting the species (considered as supplementary variables) in the factorial space, is not a valid method, because the axes correspond to linear combinations of the environmental variables based on exact statistical properties, but without physical meaning. Taking the space of the species instead, and projecting the ecological factors in it, is also inappropriate because correlations do not provide a meaningful way of statistically modelling the relationships between species and ecological variables. CA performed on contingency tables (classes of biological against classes of

ecological variables) avoids the limitation of linearity of the relations, but the meaning of the factorial space so obtained by mixing biological and physical variables remains unclear. The classification and ordination obtained using the phytosociological method, on the other hand, are meant to represent proximities among species as characterised by their responses to the intensity of the ecological variables. In our test data, for instance, some species assemblages were easy to categorise: cladocerans, echinoderm larvae and pteropods belong to the typical coastal niche, characterised by high temperature and low salinity, density and fluorescence. Ter Braak's (1986, 1987) canonical correspondence analysis (CCA) would be another correct way of representing species-environmental relationships. The question to be answered by CCA is slightly different, however, since in CCA one is looking for the species' CA ordination axes which are maximally correlated to the space of the environmental variables; biplots can be produced to represent the species-environment correlations.

The biological space corresponds to the concept of biological community. Instead of defining assemblages based on pairwise abundance distances (PCA), or through the comparison of their frequencies (CA), the distance between 2 species is estimated as a function of their relationship with the remaining species. So a community of *Acartia clausi* (1) – *Oithona* sp. (in Fig. 13a; cf. the variations in Fig. 4) may be highlighted, based on co-occurrence only, even if the spatio-temporal covariation of these species in terms of their abundances is not obvious. Species associations in the biological space depend on ecological adaptations, and also on species behaviours: reproduction, mortality, generation timing, predation, mobility, etc. So, depending on the development stage, the observed distinction of copepodites 2 and 3 from the other stages and adults of *Calanus helgolandicus* emphasises strategic differences related to types of nutrition and mobility.

Determining the associations found in the biological space remains a difficult problem, however. Among others, Boucher (1984), Ibanez & Boucher (1987), Williams (1988), and Williams & Conway (1988) have emphasised the complexity of this question, due to the diversity of behaviours of marine organisms.

Mantel tests were not part of the original proposal of Brisse & Grandjouan. Using partial Mantel tests to remove the influence of the environmental variables from the classification in biological space pointed out the independence of the behaviours of some species from the spatio-temporal and environmental frameworks. The taxonomic associations found after removing the environmental influence allowed us to suggest the causes of the observed relationships.

The phytosociological method of Brisse & Grandjouan is characterised by a conceptual unity. It does not rely on using standard mathematics, as one does when applying classical ordination methods to species abundance data. Instead, its main ideas are to partition all descriptors into classes, and to estimate distances based on cumulated fidelities. It allows the ecologist to consider jointly, in a single treatment, different types of sampling (samplings in time and space, vertically, at different time scales, or with different sampling frequencies), and to define for each species a list of its preferred biological associations and a description of its ecological niche and its preferred geographic or seasonal locations. Because of these characteristics, the phytosociological method can be a powerful tool in the interpretation of marine ecological phenomena.

Acknowledgements. This study was part of the research program PNOC (Programme National d'Océanographie Côtière). The authors express their sincere gratitude to H. Brisse and G. Grandjouan for their cooperation. Cordial thanks to M. Etienne for her help in information programming.

LITERATURE CITED

- Boucher, J. (1984). Localization of zooplankton populations in the Ligurian marine front: role of ontogenic migration. *Deep Sea Res.* 31: 469–478
- Boucher, J., Ibanez, F., Prieur, L. (1987). Daily and seasonal variations in the spatial distribution of zooplankton populations in relation to the physical structure in the Ligurian Sea front. *J. mar. Res.* 45: 133–174
- Boucher, J. (1988). Space-time aspects in the dynamics of planktonic stages. In: Rothschild, B. J. (ed.) *Toward a theory on biological-physical interactions in the world oceans.* Kluwer Academic, Amsterdam, p. 203–214
- Brisse, H., Grandjouan, G. (1978). *Éléments de phytoclimatologie numérique en milieu naturel.* Actes 3^{ème} Colloque de l'Association internationale Informatique et Biosphère, Paris, 1975, p. 55–184
- Brisse, H., Grandjouan, G. (1980a). Plantes indicatrices du climat. I: Méthode de caractérisation climatique des plantes à différents niveaux d'abondance. *Bull. Soc. bot. Fr.* 127: 263–273
- Brisse, H., Grandjouan, G. (1980b). Plantes indicatrices du climat. II: Quelques exemples à différents niveaux d'abondance. *Bull. Soc. bot. Fr.* 127: 471–482
- Brisse, H., Grandjouan, G. (1981). Tendances actuelles de la biogéographie. *Mem. Soc. Biogéogr. (3^e série) II:* 113–137
- Brisse, H., Grandjouan, G., de Ruffray, P., Hoff, M. (1985). Exploitation d'une banque de données phytosociologiques. *Colloques phytosociologiques XIV,* Nancy
- Gordon, A. D. (1987). A review of hierarchical classification. *J. R. Stat. Soc. (Ser. A)* 150: 119–137
- Grandjouan, G. (1982a). Une méthode de comparaison statistique entre les répartitions des plantes et des climats. Thèse d'État, Université Louis Pasteur, Strasbourg
- Grandjouan, G. (1982b). Phytosociologie et phytoécologie deux disciplines complémentaires. *Coll. méthodes math. appl. géogr., journées des 7 et 8 Octobre, 1982.* Université de Franche Comté, Besançon

- Hutchinson, G. E. (1957). Concluding remarks. Cold Spring Harbor Symp. quant. Biol. 22: 415–427
- Ibanez, F. (1982). Programme PRETROPHOS: étude de la micro-distribution spatio-temporelle du zooplancton en relation avec les événements hydrologiques. Rapport CNEXO, contrat 81.9485: 1–35, IFREMER, Brest
- Ibanez, F. (1984). Sur la segmentation des séries chronologiques planctoniques multivariées. *Oceanologica Acta* 7: 481–492
- Ibanez, F., Boucher, J. (1987). Anisotropie des populations zooplanctoniques dans la zone frontale de la Mer Ligure. *Oceanologica Acta* 10: 205–216
- Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a numerical method. *Psychometrica* 29: 1–27
- Lance, G. N., Williams, W. T. (1967). A general theory of classificatory sorting strategies. I. Hierarchical systems. *Computer J.* 9: 373–380
- Legendre, L., Legendre, P. (1983). Numerical ecology. Developments in environmental modelling, 3. Elsevier, Amsterdam
- Legendre, L., Legendre, P. (1984a). *Écologie numérique*, 2^{ème} éd. Tome 2: La structure des données écologiques. Masson, Paris, et les Presses de l'Université du Québec
- Legendre, P. (1987). Constrained clustering. In: Legendre, L., Legendre, P. (eds.) *Developments in numerical ecology*. NATO ASI Series, Vol. G 14. Springer-Verlag, Berlin, p. 289–307
- Legendre, P., Dallot, S., Legendre, L. (1985). Succession of species within a community: chronological clustering with application to marine and freshwater zooplankton. *Am. Nat.* 125: 257–288
- Legendre, P., Fortin, M. J. (1989). Spatial pattern and ecological analysis. *Vegetatio* 80: 107–138
- Legendre, P., Legendre, V. (1984b). Postglacial dispersal of freshwater fishes in the Québec peninsula. *Can. J. Fish. Aquat. Sci.* 41: 1781–1802
- Legendre, P., Troussellier, M. (1988). Aquatic heterotrophic bacteria: modeling in the presence of spatial autocorrelation. *Limnol. Oceanogr.* 33: 1055–1067
- Malthus, T. R. (1798). An essay on the principle of population as it affects the future improvement of society with remarks on the speculations of Mr. Godwin, Mr. Condorcet and other writers. Johnson, London
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209–220
- Smouse, P. E., Long, C. J., Sokal, R. R. (1986). Multiple regression and correlations extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35: 627–632
- ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167–1179
- ter Braak, C. J. F. (1987). Ordination. Chap. 5. In: Jongman, R. H. G., ter Braak, C. J. F., van Tongeren, O. F. R. (eds.) *Data analysis in community and landscape ecology*. PUDOC, Wageningen, p. 91–173
- Williams, R. (1988). Spatial heterogeneity and niche differentiation in oceanic zooplankton. *Hydrobiologia* 167/168: 151–159
- Williams R., Conway, D. V. P. (1988). Vertical distribution and seasonal numerical abundance of the Calanidae in oceanic waters to the south-west of the British Isles. *Hydrobiologia* 167/168: 259–266

This article was submitted to the editor

Manuscript first received: July 10, 1992

Revised version accepted: December 28, 1992