

# Le cycle annuel du zooplancton à Pointe-Noire (RP Congo) Description mathématique

Zooplankton  
Cycle annuel  
Traitement mathématique  
Codage  
Analyse factorielle  
Ordination  
Hydrologie  
Zooplankton  
Annual cycle  
Data processing  
Coding  
Factorial Analysis  
Ordination  
Hydrology

A. Dessier <sup>a</sup>, A. Laurec <sup>b</sup>

<sup>a</sup> Orstom, Centre Océanologique de Bretagne, 29273 Brest.

<sup>b</sup> Centre Océanologique de Bretagne, 29273 Brest.

Reçu le 7/11/77, révisé le 28/3/78, accepté le 6/4/78.

## RÉSUMÉ

6 années de récolte de zooplancton, à partir d'un wharf minéralier, ont été analysées sur une base hebdomadaire. L'étude porte sur une description d'un cycle annuel moyen, sur ses liens avec les différents taxons, enfin sur les particularités annuelles, l'accent étant constamment mis sur les relations avec l'hydrologie. Elle comporte un volet semi-quantitatif où l'information est traduite en variables logiques (exemple : absence, présence en faible densité, grande abondance d'une espèce), et un volet proprement quantitatif où les effectifs des taxons dans les prélèvements sont utilisés après transformation. Elle fait appel à des analyses factorielles et comporte à ce propos, des remarques méthodologiques. Les analyses révèlent un cycle saisonnier relativement simple, directement relié aux saisons hydrologiques. Le lien se retrouve, de façon très étroite, au niveau des singularités annuelles pour le zooplancton et l'hydrologie.

*Oceanol. Acta*, 1, 3, 1978, 285-304.

## ABSTRACT

### Zooplankton annual cycle at Pointe-Noire (RP Congo) mathematical description

As part of a general oceanological programme, zooplankton samples were collected over a period of 6 years, at least weekly, from a wharf off Pointe-Noire (4°49.2'S; 11°50.3'E). This paper provides a description of the zooplankton seasonal cycle, with special emphasis on the relationships with hydrology. Considering an average year, on a weekly basis, its first aim is to summarize the main features of the similarities and dissimilarities between the weeks of this average cycle. The second aim is to relate these features to different taxa. The paper concludes with a study of annual peculiarities. Mathematical data processing is used to achieve these ends. In addition to its concern with plankton, the paper has a methodological aspect, related to the use of mathematical data processing. The first stage involves the definition of mathematical variables associated with zooplanktonic taxa. Two types of variable are used: logical (or binary) and quantitative. The first information provided by a taxon is simply related to its presence or its absence. A logical variable may be associated with the latter. Apart from this, variations exist with regard to the abundance of zooplankton. For each species a threshold is defined, associated with the value of the abundance histogram, samples implying absence being

excluded. Three basic situations are considered for each species: absence; presence, but in small quantities; abundance. A logical variable is associated with each situation, taking the value 1 if the logical condition is satisfied for the sample, the value 0 if not. This basic schema cannot always be used, and a more detailed definition of logical variables is provided in Table 1. The study using logical variables will be termed semi-quantitative.

Use of logical variables makes it possible to take account of the peculiarity of the information associated with the presence or absence of a species. On the other hand, quantitative variations are barely taken into account. This makes it necessary to conduct, as a counterpart, a quantitative study, with the use of quantitative variables directly derived from the abundance of a species in a sample by an eventual transformation. Such a transformation may have two different objectives. The first is to obtain more significant numbers, from an ecological point of view: if one attempts to relate zooplankton variations to environmental conditions, the relative importance of variations within high levels of abundance must be reduced. On the other hand, from a statistical point of view, the estimation of dissimilarities between weeks of an average year is generally biased. A transformation must be used, so that the bias will not be larger for dissimilarities between rich weeks. Taking these two points of view into account, a transformation by  $\text{Log}^2(x+1)$  is used.

In the semi-quantitative study, a table directly related to a contingency table is defined. The logical variables are associated with the columns, the rows with the 52 weeks of a hypothetical average year. All the samples have been gathered, with a weighting designed to give the same influence to all individual years (the sampling step has changed from year to year). Rows corresponding to the weeks of the six individual years have been added to this basic table. In the quantitative approach, a table has also been defined with 52 rows, and columns associated with the variables. The value of a variable for a given week is provided by a weighted average of the values for the individual samples corresponding to this week. The weighting is defined as before, so that within each week of the average year, each individual year gets the same influence.

The tables defined above have been subjected to factorial (or inertia) analysis, techniques which have been chosen in order to achieve the goals set out at the beginning of the study, and which are considered as metric ordination techniques. If one considers the set of 52 weeks of the average year, a distance being chosen which combines the influences of the different variables, ordination techniques will provide configurations of low dimensionality, giving the best summary of the set of mutual distances between weeks. This corresponds to the first aim of the study. The use of metric ordination techniques in their more general form, principal co-ordinate analysis (Gower, 1966), or non-metric ordination techniques would make it difficult to achieve the second aim. Consequently, the studies use only techniques directly derived from principal component analysis, including correspondence analysis (reciprocal averaging). Three factorial analyses have been performed. The semi-quantitative approach has led to a correspondence analyses. In the quantitative approach, two principal components analyses have been performed, the first associated with the correlation matrix between species, and the second with the variance-covariance matrix. Following Chardy *et al.* (1976) special attention is paid to the ecological significance of the choice of a given technique. These methodological points are detailed in several appendices: appendix A for the use of a given ordination technique, appendix B for a number of remarks on statistical inference problems and appendix C for relations with other studies, using other techniques, such as time series analyses.

In order to achieve the last objective of this paper, namely the study of annual peculiarities, the weeks of the individual years have been associated with the average year during the correspondence analyses. They are considered as witness points, with zero weight (Benzecri *et al.*, 1973). Study of the structure of the average year shows that the results of the three analyses are very consistent. The classical hydrological seasons (Berrit, 1958) appear in a very concise and clear manner within the first axes. The first one alone provides the main features, being closely related to water temperature, which must in fact be considered as indicating the variations of a complex factor. Transitional situations, associated with sharp variations of water temperature, appear on the second axis. This is especially clear for mid-September (37th week in Figures 2 and 6). The relationships between these structures and individual species are indicated by the contributions of the variables to the axes and by dual configurations (see, for instance, Figures 2 *bis* and 6 *bis*). Annual peculiarities may be revealed by projecting the witness points on the axes, an example being given by Figure 4. Such anomalies are of three types: shifting in the appearance of the phenomena; amplitude modulation; disappearance of some

phenomena. These depend on the phenomenon; for instance, when an original fauna appears in September, it appears almost always for the same week.

Any zooplankton anomaly may be related to a hydrological phenomenon. This is obvious for the first axis in relationship to water temperature (Fig. 5) and is also true for the second. For instance, in the September transitional situation, whenever the original fauna appears it is always associated with a complete vertical mixing of the water.

If the existence of a relationship between hydrology and zooplankton is trivial, the closeness of this relationship appears very interesting. Year after year, as may be observed from the raw data, the same hydrological conditions leads to the appearance of the same fauna. Although, it is difficult to obtain significant conclusions from only single-point sampling, this may well be related to the classical problem of the stability of planktonic ecosystems. On the other hand, the study suggests that zooplankton studies could provide a powerful tool in hydrology.

*Oceanol. Acta*, 1, 3, 1978, 285-304.

## INTRODUCTION

Un programme pluriannuel d'observation du milieu marin à partir d'une station côtière, située à Pointe-Noire (République Populaire du Congo), a été défini par l'Orstom en 1969. Les données recueillies couvrent le zooplancton et un ensemble de paramètres hydrologiques et hydrobiologiques. Le but de la présente étude est d'entamer l'analyse de ces données.

La région étant relativement mal connue, la première étape doit consister en une description des phénomènes essentiels. L'étude, consacrée au zooplancton, revêtira donc un aspect essentiellement descriptif. Le fait majeur est l'existence d'un cycle saisonnier marqué, lié à l'hydroclimat. L'hydroclimat de cette région a primitivement été étudié par Berrit (1958); les données hydrologiques recueillies à la station côtière depuis 1969 n'ont, pour l'instant, fait l'objet que d'une publication partielle de la part de Dufour et Merle (1972) et nous reprendrons la nomenclature de Berrit concernant les saisons marines : Grande Saison Chaude (GSC), Grande Saison Froide (GSF), Petite Saison Chaude (PSC) et Petite Saison Froide (PSF). Nous avons représenté (fig. 1) l'évolution de la température en surface et au fond au cours de l'année moyenne 1969-1974. Les limites

temporelles des saisons hydrologiques sont données à titre indicatif. Comme l'indique son titre, cet article vise donc à une caractérisation du cycle annuel du zooplancton.

Dans un souci de synthèse nous en sommes tenus à l'aspect synécologique des phénomènes. Par ailleurs, nous n'avons pas prétendu, dans ce premier temps, aborder la dynamique des peuplements.

L'étude couvrant un grand nombre de taxons, il est indispensable de faire appel à des méthodes d'analyse multivariable. Les techniques utilisées sont essentiellement des analyses factorielles, dites encore analyses d'inertie. On n'attend pas de ces méthodes qu'elles révèlent l'existence d'un cycle saisonnier, ce fait est connu. On espère en obtenir une présentation aussi synthétique et aussi fidèle que possible des caractéristiques essentielles du cycle, de ses liens avec les différents taxons. Si les techniques évoquées sont désormais d'un usage courant, leur utilisation n'est pas toujours assortie d'une réflexion conjuguée, écologique et mathématique. Un effort sera accompli dans cet esprit afin de choisir les variantes les mieux adaptées, de tirer un profit maximal du traitement.

L'étude présentée ne correspond qu'à un premier dépouillement. Ces études en cours visent à un examen

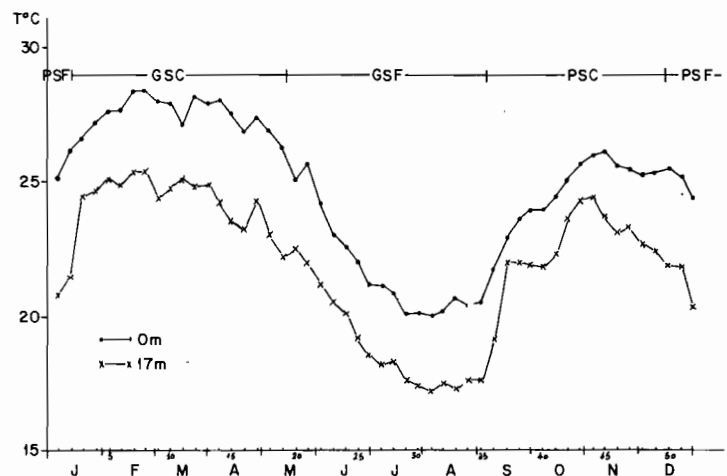


Figure 1

Moyennes hebdomadaires des températures en surface et au fond (17 m) pour la période 1969-1974. Les limites des saisons marines sont données à titre indicatif.

Weekly averages for water temperature, surface and bottom. The limits of the hydrological seasons (GSC = Long Warm Season; GSF = Long Cold Season; PSC = Short Warm Season; PSF = Short Cold Season) are given as mere indications.

sous un angle plus dynamique en intégrant l'ensemble des variables mesurées. Par ailleurs l'environnement biogéographique de la station, et notamment les variations saisonnières de cet environnement devront être analysés. Ces points seront abordés dans des publications ultérieures.

## MATÉRIEL

Le matériel utilisé dans cette étude provient d'une station côtière, située à Pointe-Noire par 4°49,2'S et 11°50,3'E, à l'extrémité d'un wharf minéralier, sur un fond moyen de 17 m. Cette station a été régulièrement visitée depuis avril 1969. La fréquence des prélèvements de zooplancton, dont seuls ceux antérieurs à janvier 1975 ont été dépouillés, a subi diverses modifications : la récolte fut bihebdomadaire d'avril 1969 à mai 1970, hebdomadaire de juin 1970 à septembre 1973, quotidienne d'octobre 1973 à octobre 1974; elle est hebdomadaire depuis novembre 1974.

Les pêches planctoniques, soit trois traits verticaux fond-surface successifs à l'aide d'un filet WP<sub>2</sub> standard, ont été effectuées le matin entre 7 et 8 h. Le volume filtré théorique, estimé d'après le parcours linéaire du filet était voisin de 4 m<sup>3</sup>. Il est vraisemblable que le volume réel a souvent été supérieur à cette valeur, du fait de courants plus ou moins forts, animant surtout les couches superficielles (jusque vers 5-10 m), et faisant dériver le filet. Les espèces de copépodes, et les principaux taxons, ont été dénombrés sur une partie aliquote d'un des prélèvements de chacune des stations. Pour la période allant d'avril 1969 à décembre 1974, 601 échantillons au total ont été examinés.

Outre les prélèvements du zooplancton, diverses mesures concernant l'hydrologie, la teneur en phosphates et la production primaire ont aussi été effectuées à différents niveaux. On en trouvera des résultats partiels dans Dufour et Merle (1972), Guillermet et Le Bouteiller (1972).

La liste des espèces et taxons divers que nous avons retenus pour l'analyse est donnée dans le tableau 1. Nous avons éliminé les espèces de copépodes : 26, (*Candacia* spp), présente une seule fois, 39 (*Oncaea minuta*) et 41 (*Oncaea dentipes*) dont l'identification était incertaine, ainsi que l'espèce 52, non identifiée. L'espèce 24 (*Oithona* sp.) est vraisemblablement nouvelle; elle sera décrite ultérieurement. Les *Lucifer* appartiennent à peu près certainement à une seule espèce, *L. faxonii*. Les *Sagitta* spp. juvéniles, non identifiables, sont vraisemblablement pour la plupart des *S. friderici tenuis*, espèce de chaetognathe largement prédominante dans les eaux côtières.

## MÉTHODES D'ANALYSE MULTIVARIABLES

### But de l'étude et choix des méthodes

Une analyse mathématique doit correspondre à un but écologique défini. Ce but a été esquissé dans

l'introduction. Il sera précisé. Dans un second temps nous indiquerons les techniques utilisées, avant de présenter quelques réflexions sur leur adéquation.

### But écologique de l'étude

Le but poursuivi est limité : description du cycle annuel du zooplancton. Par cycle annuel nous entendons d'abord un cycle moyen. L'étude des particularités de chaque année individuelle n'interviendra que dans un second temps. L'étude sera effectuée sur une base hebdomadaire. Décrire le cycle saisonnier c'est alors rendre compte des ressemblances entre semaines, et préciser les liens du cycle avec la présence et l'absence des différents taxons, avec leurs variations d'abondance. Les taxons sont essentiellement étudiés en tant que révélateurs des conditions de l'environnement. (Nous aurions volontiers parlé d'indicateur, si ce terme n'avait pas été utilisé dans un sens restreint, celui d'identificateur d'une masse d'eau.) Ce n'est qu'à titre de compléments qu'une étude sera présentée, qui donnera plus d'importance aux taxons numériquement importants, susceptibles de jouer un rôle prépondérant dans l'économie de l'écosystème.

### Analyses factorielles ou analyses d'inertie.

Soit un tableau X (I, J), où les I lignes sont associées à des observations, les J colonnes à autant de variables. Les analyses factorielles seront envisagées comme des techniques de description d'un tel tableau, selon le point de vue exposé entre autres par Benzecri *et al.* (1973), Cailliez et Pages (1976), Lebart *et al.* (1977). En tant que telles elles ne requièrent (Prentice, 1977) aucune hypothèse de multilinéarité. Il n'est pas non plus nécessaire d'admettre un modèle stochastique où les observations constituent autant de réalisations indépendantes d'une même loi.

Si l'on met l'accent sur la structure de l'ensemble des observations on définira d'abord une mesure mathématique de dissemblance entre observations, qui sera dans cet article une distance. Des masses seront données aux différentes observations. Les observations de masse nulle, qui ne joueront pas sur les résultats des analyses, constituent des observations dites témoins. On cherchera des configurations de dimension réduite rendant compte au mieux du réseau de dissemblance entre observations (les observations témoins étant exclues à ce niveau) tel qu'il est défini par la matrice des distances. C'est le propos des techniques d'ordination métriques. Les configurations recherchées seront fournies par des analyses d'inertie, pratiquées à partir du barycentre des observations, en utilisant la propriété de représentation optimale (Lebart *et al.*, 1977). On dégage par ces analyses des axes dits factoriels. L'importance d'un axe est mesurée par le pourcentage d'inertie qu'il est dit expliquer. Deux par deux, les axes définissent des plans sur lesquels on projette les observations. On obtient ainsi les configurations observations. Sur ces plans, on pourra projeter les observations témoins, même si elles n'ont pas influé sur la détermination des axes factoriels.

On pourra ne pas se contenter de caractériser de façon simple la structure de l'ensemble des observations, pour tenter de préciser les liens des variables avec cette

Tableau 1

Liste des taxons, code d'identification, et définition des variables logiques associées.

List of taxa, definition of the associated logical variables.

Copépodes	Codage logique		
	1	2	3
1 <i>Nannocalanus minor</i>	0	0 <	
2 <i>Calanoides carinatus</i>	0	0 <	
3 <i>Neocalanus gracilis</i>	0	0 <	
4 <i>Undinula vulgaris</i>	0	0 <	
5 <i>Eucalanus pileatus</i>	< 100	{ 100- 400 }	> 400
6 <i>Eucalanus monachus</i>	0	0 <	
7 <i>Acrocalanus andersoni</i>	0	0 <	
8 <i>Paracalanus scotti</i>	< 700	{ 700- 12 000 }	> 12 000
9 <i>Paracalanus parvus</i>	< 450	{ 450- 4 000 }	> 4 000
10 <i>Paracalanus aculeatus</i>	0	0 < < 100 >	100
11 <i>Calocalanus clausi</i>	0	0 <	
12 <i>Mecynocera clausi</i>	0	0 <	
13 <i>Clausocalanus</i> spp.	0	0 <	
14 <i>Euchaeta marina</i>	0	0 <	
15 <i>Euchaeta paraconcinna</i>	0	0 <	
16 <i>Euchaeta</i> spp. juvéniles	0	0 <	
17 <i>Diaxis hibernica</i>	0	0 < < 30 >	30
18 <i>Centropages furcatus</i>	0	0 < < 170 >	170
19 <i>Centropages chierchiae</i>	0	0 < < 300 >	300
20 <i>Pseudodiaptomus serricaudatus</i>	< 350	{ 350 1 000 }	> 1 000
21 <i>Temora stylifera</i>	0	0 < < 100 >	100
22 <i>Temora turbinata</i>	< 25	{ 25- 700 }	> 700
23 <i>Temoropia mayumbaensis</i>	0	0 <	
24 <i>Lucicutia flavicornis</i>	0	0 <	
25 <i>Lucicutia clausi</i>	0	0 <	
27 <i>Pontella gabonensis</i>	0	0 <	
28 <i>Labidocera scotti</i>	0	0 < < 25 >	25
29 <i>Acartia plumosa</i>	< 450	{ 450 1 750 }	> 1 750
30 <i>Acartia danae</i>	0	0 <	
31 <i>Oithona nana</i>	< 1 000	{ 1 000- 11 500 }	> 11 500
32 <i>Oithona plumifera</i>	0	0 <	
33 <i>Oithona simplex</i>	0	0 < < 25 >	25
34 <i>Oithona</i> sp.	0	0 < < 50 >	50
35 <i>Microsetella</i> spp.	0	0 <	
36 <i>Macrosetella gracilis</i>	0	0 <	
37 <i>Euterpina acutifrons</i>	< 3 000	3 000 <	
38 <i>Clytemnestra</i> sp.	0	0 <	
40 <i>Oncaea curta</i>	0	0 < < 65 >	65
42 <i>Oncaea venusta</i>	0	0 <	
43 <i>Oncaea mediterranea</i>	0	0 <	
44 <i>Oncaea media</i>	0	0 <	
45 <i>Lubbockia squillimana</i>	0	0 <	

Copépodes	Codage logique		
	1	2	3
46 <i>Corycaeus speciosus</i>	0	0 <	
47 <i>Onychocorycaeus giesbrechti</i>	0	0 < < 170 >	170
48 <i>Ditrichocorycaeus africanus</i>	< 1 500	> 1 500	
49 <i>Corycella gracilis</i>	0	0 <	
50 <i>Corycella rostrata</i>	0	0 <	
51 <i>Saphirella</i> sp.	< 50	{ 50- 600 }	> 600
Divers			
53 <i>Méduses</i>	< 50	{ 50- 600 }	> 600
54 <i>Siphonophores</i>	< 25	{ 25- 100 }	> 100
55 <i>Cténophores</i>	< 25	{ 25- 100 }	> 100
56 <i>Annélides</i>	< 25	{ 25- 200 }	> 200
57 <i>Total Chaetognathes</i>	< 1 200	> 1 200	
58 <i>Larves actinotroques</i>	0	0 <	
59 <i>Penilia</i> sp.	< 25	{ 25 1 000 }	> 1 000
60 <i>Evadne</i> sp.	0	0 <	
61 <i>Ostracodes</i>	0	0 <	
62 <i>Mysidacés</i>	0	0 < < 100 >	100
63 <i>Lucifer faxonii zoés</i>	0	0 < < 300 >	300
64 <i>Lucifer faxonii mysis</i>	0	0 < < 300 >	300
65 <i>Lucifer faxonii mastigopus</i>	0	0 <	
66 <i>Lucifer faxonii adultes</i>	0	0 < < 25 >	25
67 <i>Total Lucifer faxonii</i>	0	0 < < 200 >	200
68 <i>Nauplii (copépodes)</i>	0	0 <	
69 <i>Larves de décapodes</i>	< 100	100 >	
70 <i>Zoés de branchyours</i>	< 50	50 <	
71 <i>Mégalopes de branchyours</i>	0	0 <	
72 <i>Larves de stomatopodes</i>	0	0 <	
73 <i>Appendiculaires</i>	< 700	{ 700- 3 400 }	> 3 400
74 <i>Doliolés</i>	< 25	25- 300	> 300
75 <i>Œufs de poisson</i>	0	0 <	
76 <i>Larves de poisson</i>	0	0 <	
77 <i>Sagitta enflata</i>	0	0 <	
78 <i>Sagitta tenuis-friderici</i>	< 100	{ 100- 400 }	> 400
79 <i>Sagitta hispida</i>	0	0 <	
80 <i>Sagitta</i> sp. juvéniles	< 300	300 <	
81 <i>Krohnitta pacifica</i>	0	0 <	

structure. Ce sera spécialement facile lorsque l'on utilisera des structures mathématiques simples (structures euclidiennes), qui permettent de jouer sur la notion de dualité. Ceci est spécialement fructueux lorsque la distance d'une entre deux observations  $i_1$  et  $i_2$  est de la forme

$$d^2_{i_1, i_2} = \sum_j pc_j (x_{i_1, j} - x_{i_2, j})^2,$$

$pc_j$  étant la masse attribuée à la variable  $j$ . Ceci correspond à une structure euclidienne simple où le produit scalaire des lignes  $i_1$  et  $i_2$  est égal à  $\sum_j pc_j x_{i_1, j} x_{i_2, j}$ .

Les analyses factorielles les plus classiques peuvent être ramenées à ces structures simples en jouant sur la

pondération et éventuellement des transformations préliminaires du tableau (Chardy *et al.*, 1976).

Dans ces structures une analyse est effectuée à partir du barycentre des observations si les colonnes sont centrées, ayant une moyenne nulle. Dans les structures simples évoquées, en généralisant des notions d'un usage courant dans le cas particulier de l'analyse des correspondances, on définira d'abord la contribution relative des variables aux axes. Cette contribution quantifie la « responsabilité » d'une variable vis-à-vis d'un axe factoriel. Sans aborder les notions de contribution des axes aux variables, qui ne seront pas utilisées dans cet article, on peut encore définir des facteurs, associés aux axes factoriels. Ce sont des variables « abstraites », définies sur l'ensemble des observations, par rapport auxquelles on pourra décomposer les variables originelles.

Cette décomposition permet de définir des coordonnées des variables par rapport aux facteurs. Ces coordonnées seront utilisées pour bâtir des configurations variables, dites duales des configurations observations. Il existe une ambiguïté quant à l'utilisation de ces configurations. On peut tendre à les utiliser comme rendant compte des proximités entre variables. Ce peut être justifié, ce ne l'est pas nécessairement comme le précise l'annexe A. Les configurations entre variables traduisent dans l'ensemble des liens entre les variables et les facteurs, plus que des liens directs entre variables. On peut vouloir encore rechercher des associations particulières entre observations et variables en examinant simultanément les configurations duales. Cet examen peut être plus ou moins fructueux selon les variantes des analyses factorielles. Enfin les configurations variables peuvent dans certains cas offrir une représentation graphique des contributions. Ces points doivent être examinés pour chaque variante, comme le fait l'annexe A.

#### *Adéquation des outils mathématiques utilisés*

L'approche dite descriptive des analyses factorielles apporte un outil précieux, si l'on se reporte aux préoccupations écologiques annoncées. On se ramènera à un tableau bidimensionnel où les 1=52 lignes correspondront aux semaines d'un cycle complet, où les colonnes (variables) seront associées aux différents taxons. Une distance sera définie pour quantifier les dissemblances entre semaines. Ce sera une simple distance euclidienne avec pondération, et éventuellement transformation préliminaire simple du tableau. On cherchera ensuite les configurations de dimension réduite respectant au mieux le réseau de dissemblances entre semaines défini par la matrice des distances. On fera ainsi appel à une analyse d'inertie à partir du barycentre des semaines, en jouant sur la notion de représentation optimale. En se limitant aux premiers axes de l'analyse, pour s'en tenir aux phénomènes essentiels, on pourra étudier les particularités des années individuelles. Pour ce faire, on adjoindra aux 52 lignes définissant dans le tableau l'année moyenne, des lignes supplémentaires correspondant aux semaines des années individuelles. Ce seront des éléments témoins. Le lien des structures ainsi dégagées avec les taxons étudiés sera précisé par l'examen des contributions et des structures duales.

S'en tenir aux structures simples évoquées interdit le recours à des mesures de la ressemblance ou de la dissemblance qui pourraient être utiles et utilisées dans le cadre d'analyses en coordonnées principales (Gower, 1966). On ne fera pas non plus appel aux techniques d'ordination non métriques, qui présentent pourtant un intérêt certain (Prentice, 1977; Fasham, 1977). Ces techniques rendent plus difficiles l'étude du lien des variables avec les structures de l'ensemble des observations. Le souci d'une étude fine de ce lien conduit à limiter les procédés de mesure des dissemblances entre semaines aux distances euclidiennes compliquées d'une simple pondération.

#### **Constitution du tableau bidimensionnel**

##### *Définition des colonnes : codage de l'information*

L'information apportée par les différents taxons doit être

codée numériquement : selon une expression empruntée à Reys (1976), on doit associer aux taxons des descripteurs numériques qui constitueront les colonnes du tableau. Ces variables seront logiques ou binaires lorsque l'on se ramènera à des études en présence absence, ou de façon plus générale à des conditions satisfaites ou non. Lorsque l'on utilisera les effectifs, après une éventuelle transformation, on définira des variables quantitatives.

L'information apportée par un taxon peut être d'une double nature : soit une présence ou une absence, soit une variation d'abondance. A l'absence est associée une variable logique. Par souci d'homogénéité on peut également refléter les variations d'abondance par des variables logiques. Un seuil sera fixé, correspondant à la valeur médiane de l'histogramme de l'abondance d'un taxon, les absences étant exclues. Ceci a pour but de créer des classes sensiblement équiprobables, selon la technique de Cros et Durbec (*in* Benzecri *et al.*, 1973). Ce schéma idéal n'a pas toujours pu être respecté, ainsi que le révèle le tableau 1, où les seuils choisis ont été reportés. Certains taxons sont trop rarement présents (*cf.* taxon 9), ou présentent des variations d'abondance trop peu soumises au cycle saisonnier (taxon 43) pour que les fluctuations d'abondance soient prises en compte. Deux variables logiques sont alors créées, correspondant à la présence et à l'absence. D'autres taxons (taxon 5) sont trop rarement absents pour qu'une variable logique particulière soit associée à cet événement. Trois classes d'abondance, sensiblement équiprobables sont créées. Si les fluctuations d'abondance ne paraissent pas nettement liées au cycle annuel, deux classes d'abondances apparaîtront (*cf.* taxon 37). Le choix du codage passe par l'examen et de l'histogramme de chaque taxon, et de sa courbe d'évolution annuelle. Tout codage logique ne traduit qu'imparfaitement les fluctuations d'abondance. D'une part, une discontinuité est artificiellement créée par l'introduction de seuils, d'autre part les fluctuations entre deux seuils ne sont pas prises en compte.

On peut alors en venir à des méthodes quantitatives et associer à chaque taxon une variable donnée par son effectif, éventuellement transformé, dans un prélèvement. Une transformation répond à un double objectif : définition d'une échelle de variation qui soit écologiquement plus significative, et satisfaction d'impératifs statistiques. En ce qui concerne le second point, comme le détaille l'annexe B, les lignes du tableau analysé correspondent à des périodes précises. L'abondance moyenne durant ces périodes est estimée à l'aide des échantillons collectés. Une erreur d'échantillonnage existe, qui comporte une variance allant croissant selon l'importance des abondances réelles. Lorsque l'on estime les distances entre semaines, cette variance introduit un biais qui sera plus grand pour les périodes de forte abondance. Pour éviter ou plus modestement limiter ce phénomène, on peut faire appel à une transformation visant à stabiliser les variances d'échantillonnage, c'est-à-dire les rendre indépendantes des niveaux d'abondance. On remarquera que le but n'est pas d'homogénéiser la variance des différentes colonnes, mais d'éviter une distorsion dans l'estimation des distances entre semaines. Suivant en cela Frontier (1973), nous avons utilisé une



transformation  $\text{Log}^\alpha(x+1)$ , avec  $\alpha=2$ . Le biais dans l'estimation des distances entre semaines ne dépend plus, en théorie, de la richesse des périodes intéressées.

Une transformation peut encore viser, si l'on tente d'utiliser les taxons comme des révélateurs de l'environnement, à déterminer une échelle de variation qui soit la plus significative possible. Dans cet esprit on sera tenté d'accorder plus de signification à un même écart absolu dans les niveaux d'abondance les plus faibles. On aura donc recours à des transformations écrasant les variations dans les forts effectifs, telle  $\text{Log}^\alpha(x+1)$  d'autant plus écrasante que  $\alpha$  est faible.

Quel que soit le critère retenu, il n'existe pas d'argument définitif et universel pour déterminer la valeur de  $\alpha$ . Si l'on se souvient que l'étude quantitative entreprise le fut en contrepoint de celle utilisant un codage logique, on comprendra que nous ayons choisi  $\alpha=2$ , transformation relativement peu écrasante.

Si l'approche semi-quantitative ne traduit qu'imparfaitement les variations d'abondance, l'approche quantitative respectera mal l'importance du saut de l'absence à la présence. Même en utilisant des transformations écrasantes, on ne gardera pas à ce type d'événement l'importance qu'il mérite.

#### Définition des lignes du tableau

L'étude est centrée sur une définition du cycle moyen. Raisonnant sur une base hebdomadaire, nous avons donc cherché à définir un tableau regroupant 52 lignes associées aux semaines d'une année moyenne. Pour définir cette année moyenne, nous avons utilisé tous les prélèvements, chacun étant rapporté à la semaine adéquate. Selon l'année, le nombre de prélèvements dans chaque semaine varie. Pour éviter que les années où la maille d'étude fut très serrée revêtent une importance prépondérante, nous avons utilisé des moyennes pondérées. Si pour l'année  $k$ ,  $N_{i,k}$  prélèvements furent effectués dans la semaine  $i$ , ceux-ci recevront un poids  $1/N_{i,k}$ .

Si l'on utilise le codage logique, on aboutit ainsi à un tableau à 52 lignes et 194 colonnes, puisque 194 variables furent définies. A ce tableau de base furent adjointes des lignes supplémentaires, correspondant aux années individuelles. Pour chaque année individuelle, toute semaine où au moins un prélèvement fut effectué engendre une ligne supplémentaire, définie par une moyenne simple des prélèvements correspondants.

Dans l'approche quantitative, avec simple transformation des effectifs, les 52 lignes correspondant à l'année moyenne furent définies de même par des moyennes pondérées. Ce tableau de base fut également complété par des lignes supplémentaires; pour limiter le volume du présent article nous ne présenterons cependant pas les résultats relatifs aux points témoins; ils ne présentent pas d'originalité marquante.

Si la définition d'une année moyenne correspond au désir de centrer l'étude sur un cycle moyen, ceci comporte un risque. On peut créer par une moyenne un peuplement théorique, intermédiaire entre des peuplements réels, mais sans existence concrète. Pour se prémunir contre le risque de mésinterprétation on peut se ramener

constamment aux données de base au niveau des interprétations. C'est ce que nous fîmes. On peut aussi envisager l'étude du tableau complet de l'ensemble des prélèvements individuels, et la comparer à celle portant sur l'année moyenne. C'est ce que nous avons également fait. Les résultats concordent. On ne présentera ainsi que l'analyse portant sur l'année moyenne, complétée par l'étude des années individuelles en tant qu'éléments supplémentaires dans l'approche semi-quantitative.

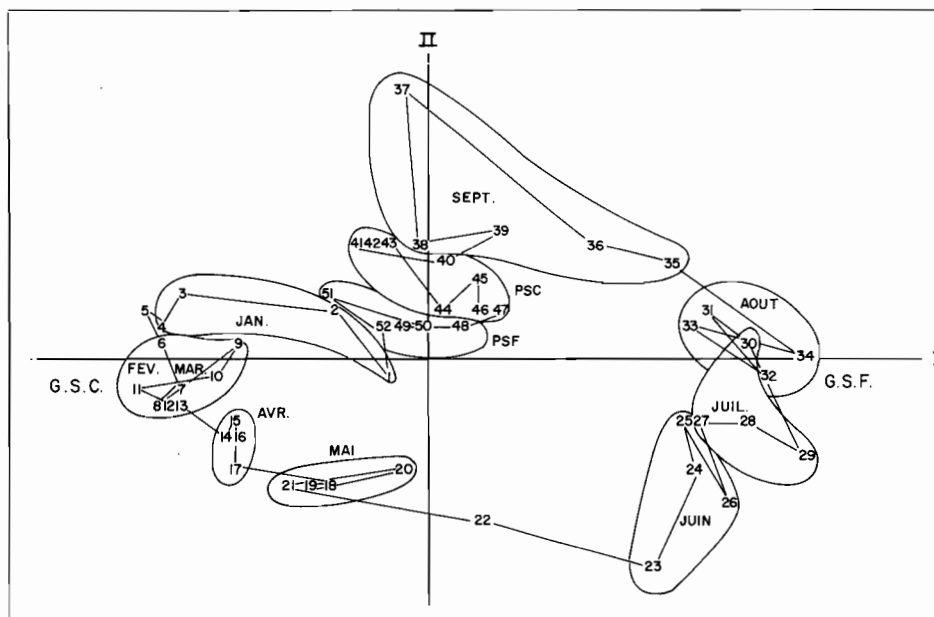
#### Étude des tableaux par des analyses d'inertie

Dans le cadre défini pour les analyses d'inertie, de nombreuses variantes peuvent être définies, au nombre desquelles figurent les classiques analyses en composantes principales et l'analyse des correspondances. Le choix d'une variante particulière doit se faire non en cherchant s'il existe une technique universellement plus puissante, mais en connaissant les implications écologiques des choix mathématiques définissant chaque variante. Au niveau de la structure des semaines ceci implique une réflexion sur la définition des distances, des masses attribuées aux différentes observations (Chardy *et al.*, 1976). Le troisième point évoqué par ces auteurs, celui du choix de l'origine, ne pose pas de problèmes, toutes les analyses d'inertie étant effectuées à partir du barycentre des semaines.

Pour comprendre le lien entre les structures dégagées dans l'ensemble des semaines et les différentes variables, et notamment bien utiliser les configurations variables, il convient de réfléchir sur la relation de ces configurations avec les contributions, sur la distance entre variables implicitement utilisée par l'analyse et sur la valeur des représentations offertes par les projections sur les premiers facteurs. Trois analyses ont été effectuées. Les réflexions qui viennent d'être suggérées constituent l'annexe A. On s'y reportera si nécessaire. Le tableau de contingence constitué à l'issue du codage logique et caractérisant une année moyenne a été soumis à une analyse des correspondances, où les lignes supplémentaires associées aux années individuelles ont été utilisées comme observations témoins. Le tableau caractérisant l'année moyenne dans l'approche quantitative a fait l'objet de deux analyses en composantes principales. La première porte sur la matrice de corrélation entre taxons. Dans ce cas tous les taxons apportent la même contribution globale à la structure de l'ensemble des semaines. La seconde porte sur la matrice de variance-covariance entre taxons. Dans ce cas (Pielou, 1969, Chardy *et al.*, 1976) les taxons numériquement abondants ont une influence prédominante. Ceci n'est pas justifié si les taxons sont des révélateurs de leur environnement. Ce peut l'être dans la mesure où l'on effectue un premier pas vers une approche plus fonctionnelle de l'étude de l'écosystème. En toute première approximation il peut être intéressant d'accorder une influence prédominante aux taxons abondants, pour voir dans quelle mesure ce changement de point de vue entraîne une modification des résultats.

Des trois analyses présentées, ce qui sépare l'analyse des correspondances après codage logique de l'analyse en composante principale sur la matrice de corrélation des

Figure 2



Figures 2 et 2 bis

Année moyenne en semi-quantitatif : plan des axes I-II issu de l'analyse des correspondances [observations = semaines (fig. 2); variables = taxons (fig. 2 bis)], codage en deux états : 1 = □ et 2 = ■ ou trois états : 1 = ○, 2 = ⊙ et 3 = ●; cf. tableau 1.

taxons, c'est essentiellement la nature de l'information retenue, la façon dont elle est traduite mathématiquement. L'utilisation de variables non réduites dans la troisième analyse correspond à une modification plus importante, sinon des outils mathématiques mis en œuvre, du moins du point de vue écologique adopté.

## RÉSULTATS

### Année moyenne en semi-quantitatif

#### Cycle saisonnier

Dans l'analyse des correspondances le pourcentage d'inertie des cinq premiers axes, respectivement (31.0, 9.4, 7.6, 5.4, 3.5) montre que les trois premiers rendent compte de l'essentiel des structures. Le plan des deux premiers axes correspond à lui seul à 40,4 % de l'inertie. On rappelle que le but de l'analyse factorielle des correspondances n'est pas, dans notre cas, de dégager des facteurs indépendants, mais de définir un espace de dimension réduite respectant au mieux les relations de proximité entre les 52 semaines, telles qu'elles ont été définies par la distance du  $\chi^2$ . On voit ainsi que le plan des axes I et II traduit une part importante de ces structures. Sans vouloir à toute force identifier chaque axe à un phénomène particulier, lorsqu'une explication apparaît, elle est évidemment bienvenue dans ce type d'analyse. Ainsi, lorsqu'on se réfère au premier axe, on constate qu'il marque l'opposition entre les grandes saisons chaudes et froides (notées respectivement GSC et GSF). Son lien avec la température est donc évident, comme le confirmera l'étude des années individuelles. Il est vraisemblable cependant que c'est un complexe de facteurs qui opère, la température n'en étant que le meilleur révélateur. Si l'on examine non plus l'axe I isolément, mais l'ensemble des premiers axes, et notamment les plans I-II (fig. 2) et I-III (fig. 3), on constate essentiellement, en dehors de l'opposition des périodes chaudes et froides, l'individualisation marquée des transitions (en septembre, janvier et à un degré

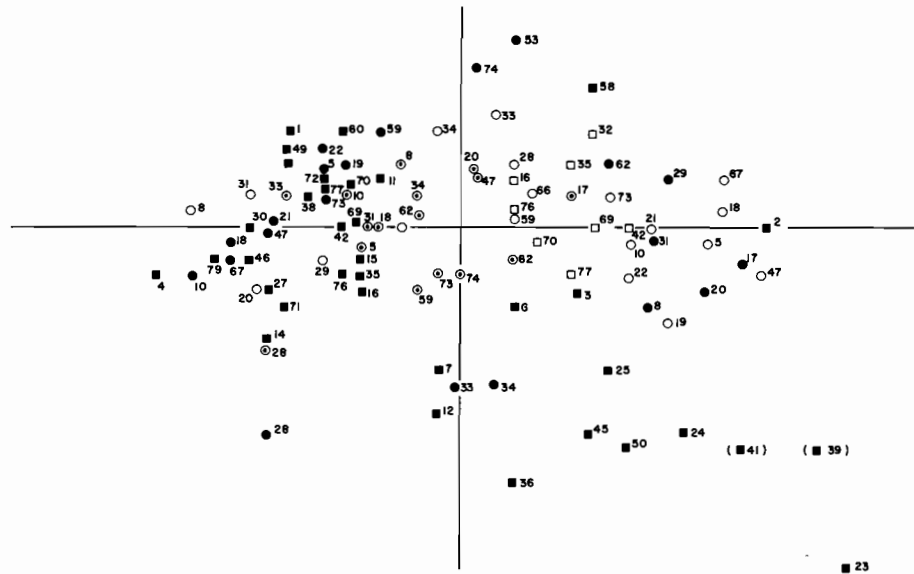
moindre début juin). La personnalité des Petites Saisons, Chaudes (PSC), et Froides (PSF) est beaucoup moins tranchée. Ces petites saisons, et ceci sera confirmé par les études quantitatives, ne correspondent pas à des peuplements originaux, mais, pour l'essentiel, à un simple moyen terme entre ceux des grandes saisons.

Ces grands traits ayant été brossés, nous allons maintenant suivre de façon plus détaillée le déroulement du cycle.

La GSC culmine en février-mars, la tendance au refroidissement manifestée durant les mois d'avril et mai relève beaucoup plus, comme le montrera l'étude des années individuelles, de refroidissements sporadiques brutaux, que d'une dérive lente des températures. L'installation de la GSF se fait très rapidement. L'isolement sur la figure 2 du début juin est lié à la persistance, en nombre progressivement décroissant de quelques espèces de GSC, comme à l'apparition en nombre réduit de quelques espèces subsuperficielles du large immédiat. Le refroidissement se poursuit progressivement désormais jusqu'à la fin du mois d'août. Il est à noter que si l'on examine la matrice des distances entre semaines, distances calculées non à partir des premiers axes, mais à partir du tableau de contingence, on voit apparaître une certaine individualisation du mois d'août, individualisation cependant trop peu marquée pour ne pas être masquée par les traits généraux du cycle. Il convient d'ailleurs pour étudier un phénomène partiel, jouant sur quelques points, d'isoler ces points pour pratiquer une analyse d'inertie séparée. Le réchauffement qui s'amorce en septembre amène une faune très caractéristique, dont l'originalité est due à l'abondance de quelques espèces ou taxons. D'un point de vue hydrologique cette faune particulière dont la singularité culmine à la 37<sup>e</sup> semaine, semble associée à l'homogénéité verticale de la couche d'eau, comme le montre l'absence de gradient vertical de densité à cette époque sur la figure 1 (cf. Merle, Guillermin, 1975). L'hypothèse la plus vraisemblable est celle du passage au wharf d'une eau originale dont la provenance demanderait à être précisée par une étude hydrologique.



Figure 2 bis



Average year, semi-quantitative study: planes defined by axes I and II. Observations = weeks (fig. 2). Variables, associated with the taxa (fig. 2 bis) two variables per taxon: 1 = □ and 2 = ■; three variables: 1 = ○, 2 = ⊙ and 3 = ● (see Table 1).

Les petites saisons qui apparaissent, ainsi que nous l'avons signalé, comme intermédiaires entre les grandes saisons, sont relativement difficiles à scinder en PSC et PSF. Dans le plan des axes II et III on voit cependant s'isoler la PSC à l'intérieur du quatrième quadrant. La mauvaise individualisation de la PSF sur l'année moyenne provient essentiellement, comme le montreront les années individuelles, d'une grande variabilité, d'une année sur l'autre, des situations de décembre et janvier. Le réchauffement qui conduira à l'installation de la GSC s'opère au cours du mois de janvier, période de transition qui présente une certaine ressemblance avec les semaines de transition menant de la GSF aux petites saisons.

#### Relations entre les différents taxons et le cycle saisonnier

Nous avons utilisé pour étudier ce rôle les contributions relatives des différentes variables aux divers axes

factoriels, et les représentations, dites duales de celles des semaines, où apparaissent les variables. Seules ces dernières seront présentées explicitement. Il est hors de propos de commenter ici l'autécologie de chacun des taxons, étude en cours pour les composants essentiels par l'un d'entre nous (A. Dessier). On remarquera cependant que l'existence de trois variables par taxon apporte une certaine finesse au niveau de l'interprétation espèce par espèce, finesse que ne procurera pas une approche quantitative. Notre commentaire se fera selon un point de vue synécologique. Pour suivre ce commentaire on se référera à la figure 2 bis, duale de la figure 2; dans cette figure 2 bis nous n'avons pas représenté les variables trop proches de l'origine.

La GSC est surtout caractérisée par l'abondance de quelques taxons (*Paracalanus aculeatus*, *Lucifer favonii*, *Centropages furcatus*) trouvant en cette saison des

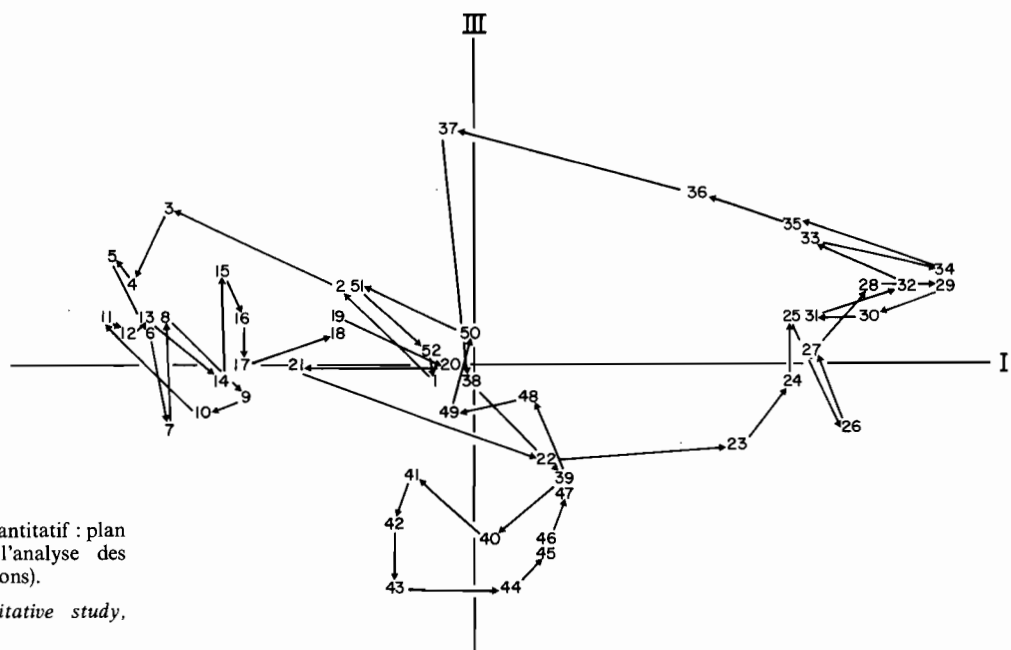


Figure 3

Année moyenne en semi-quantitatif : plan des axes I-III issu de l'analyse des correspondances (observations).

Average year, semi-quantitative study, axes I and III.

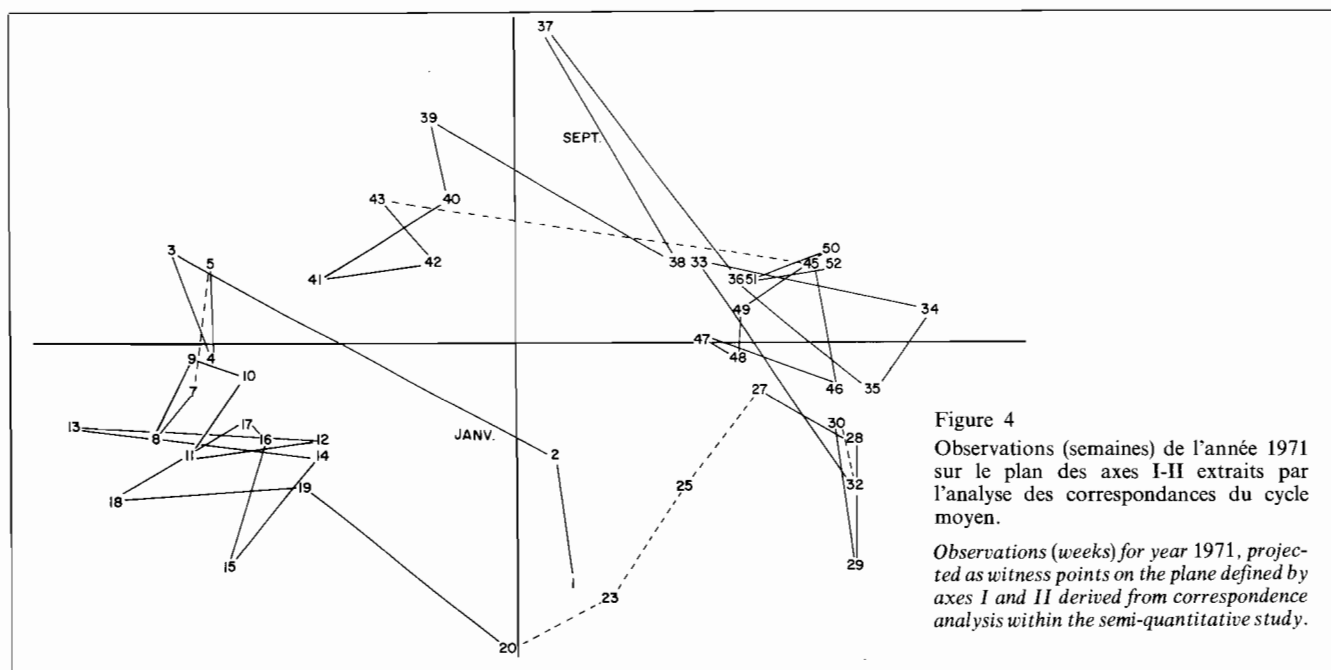


Figure 4

Observations (semaines) de l'année 1971 sur le plan des axes I-II extraits par l'analyse des correspondances du cycle moyen.

Observations (weeks) for year 1971, projected as witness points on the plane defined by axes I and II derived from correspondence analysis within the semi-quantitative study.

conditions optimales, et des présences traduisant l'influence de la faune des eaux chaudes superficielles du large (eaux guinéennes au sens de Berrit, 1958) (*Acartia danae*, *Undinula vulgaris*, *Pontella gabonensis*, *Corycaeus speciosus*, *Sagitta hispida*, etc.). D'autres espèces sont moins caractéristiques de la GSC et se retrouvent également dans les petites saisons (*Oncaea venusta*, *Eucalanus pileatus*, etc.). Les absences ou faibles abondances apparaissant comme caractéristiques de la GSC sont en fait rares et donc de contribution modeste (*Paracalanus scotti*, *Pseudodiaptomus serricaudatus*...).

De façon générale, il existe un certain nombre d'événements ne se produisant qu'à certaines époques, et donc parfaitement caractéristiques de celles-ci, mais ne survenant pas systématiquement tous les ans. Si ces événements sont rares, leur contribution aux structures des semaines sera faible, bien que dans la configuration des variables les points correspondants puissent apparaître comme extrêmement excentrés. On retrouve ainsi la distinction entre espèces structurantes et caractéristiques discutée par Blanc *et al.* (1976) à la suite de Verneaux (1973).

Certaines espèces marquent à la fois la fin de la GSC et le début de la GSF par leur maximum d'abondance (*Oithona simplex* et *Oithona* sp. par exemple) ou leur présence (*Mecynocera clausi*, *Acrocalanus andersoni*, *Macrosetella* sp., etc.).

Le début de la GSF voit l'apparition d'un certain nombre d'espèces de copépodes habituellement absentes au wharf, et dont nous avons déjà souligné l'origine sub-superficielle. L'upwelling permanent qui s'installe à cette époque (juin) est vraisemblablement responsable de l'arrivée à la côte de ces espèces du proche large (*Eucalanus monachus*, *Lucicutia clausi*, *Lucicutia flavicornis*, *Corycella rostrata*, etc.). La remarque précédente sur la distinction entre espèces caractéristiques et structurantes vaut d'ailleurs pour ces espèces qui se rattachent essentiellement à la première catégorie. On notera également, en début de GSF la persistance, en effectifs

réduits et décroissant dans le temps, d'espèces de GSC (*Eucalanus pileatus* par exemple). L'ensemble de la GSF est surtout marqué par l'abondance de quelques espèces de copépodes (*Paracalanus scotti*, *Oithona nana*, *Acartia plumosa*, *Pseudodiaptomus serricaudatus*) et la disparition, surtout en fin de saison (août), de taxons présentant leur maximum d'abondance en saison chaude (*Eucalanus pileatus*, *Centropages furcatus*, *Lucifer* sp., etc.). On s'attardera peu sur la grande abondance des mysidacés en septembre-octobre, conséquence vraisemblable d'un brassage des eaux par la houle avec mise en suspension des sédiments. La faune particulière de la mi-septembre (semaine 37) se signale par l'abondance de quelques taxons (doliolles, méduses, larves actinotroques, cladocères, etc.) et la disparition de quelques espèces de copépodes (*Oithona simplex*, *Oithona* sp., *Macrosetella* sp.).

Les petites saisons sont, pour l'essentiel, intermédiaires entre les grandes saisons comme nous l'avons annoncé. Plusieurs taxons méroplanctoniques montrent cependant une préférence pour ces petites saisons (larves de décapodes, larves actinotroques, larves de poisson) et, pour certains plus spécifiquement pendant la PSF.

Le mois de janvier, qui voit la transition vers la GSC, présente, quand la PSF est nette, certaines analogies avec le mois de septembre (abondance des cladocères et des doliolles).

#### Étude des années individuelles en semi-quantitatif

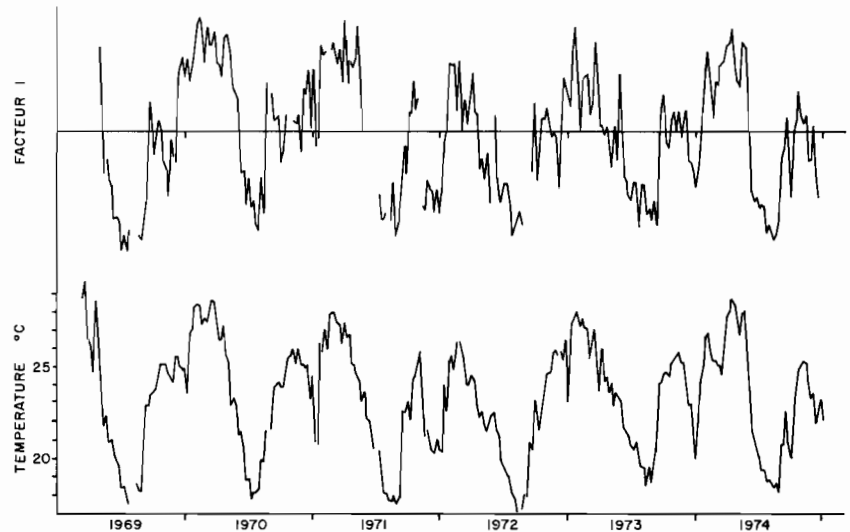
Les semaines 1 à 52 des années individuelles ont été reportées sur le plan des axes I et II extraits par l'analyse des correspondances du cycle moyen. La figure 4 correspond, à titre d'exemple, à l'année 1971. Les écarts à l'année moyenne peuvent prendre trois formes :

- la date d'installation ou de disparition de certaines conditions saisonnières peut être décalée;
- l'amplitude de certains phénomènes est plus ou moins grande;

Figure 5

Coordonnées des points témoins (observations) sur l'axe I et température moyenne (cf. texte) pour la période 1969-1974.

*Co-ordinates, on the first axis given by the correspondence analysis, of the weeks of individual years, compared with the weekly series of average water temperatures. Co-ordinates on axis I have been multiplied by  $-1$ .*



— à la limite, certains peuvent ne pas apparaître certaines années.

Les coordonnées des points témoins sur l'axe I, axe que nous avons rapproché de la température, forment une série de valeurs hebdomadaires que l'on peut comparer à celle des températures moyennes sur la tranche d'eau (estimées en effectuant une moyenne pondérée des températures aux divers niveaux de mesure, chacun recevant un poids proportionnel à l'épaisseur de la couche qu'il est censé représenter). Ces deux séries sont tracées sur la figure 5 où, pour obtenir une relation directe, et non inverse, les coordonnées sur l'axe I ont été multipliées par  $-1$ . On ne peut s'attendre à ce qu'une relation parfaitement linéaire existe. Il n'y a *a priori* aucune raison pour cela. En revanche, on remarquera que les coordonnées sur l'axe I suivent fidèlement les variations de température, paramètre de l'hydroclimat le plus facilement accessible. Cette fidélité va au-delà d'un respect du cycle saisonnier moyen, et se manifeste même lorsqu'apparaissent des anomalies. Le fait le plus spectaculaire à cet égard a trait à l'intensité de la GSC anormalement fraîche en 1972, et à la durée de la GSF anormalement courte en 1970, phénomène que les coordonnées sur le premier axe permettent bien de reconnaître. La PSF apparaît de façon très irrégulière mais parallèle, sur les deux séries. Pratiquement inexistante en décembre-janvier 1969-1970, elle est presque aussi forte en 1971-1972 qu'une GSF. Si l'on étudie les anomalies non seulement sur le premier axe, mais aussi sur le second on constate que les semaines de transition ou d'installation en juin (GSC-GSF) septembre (GSF-PSC) et janvier (PSF-GSC) peuvent ou non s'individualiser selon les années. En ce qui concerne le début du mois de juin (semaines voisines de la 23<sup>e</sup>) il faut, pour expliquer son peuplement, remonter à l'intensité et à la brutalité du refroidissement en fin de GSC. Ainsi, pour l'année 1974, ce refroidissement est particulièrement violent, et l'on voit pour cette même année le début du mois de juin (semaines 23-24) s'isoler dans la partie inférieure du quatrième quadrant. En 1972, au contraire, le refroidissement s'amorce au fond dès le mois de mars et s'étale sur près de trois mois; le début de la GSF n'est pas individualisé. On peut supposer que seul un upwelling

susceptible d'entraîner à la côte des espèces de subsurface du proche large puisse provoquer un refroidissement brutal. En janvier 1971, après que la PSF 1970 ait été limitée à un refroidissement marqué et surtout très rapide, les 2 premières semaines présentent une parenté évidente avec la situation caractéristique de l'établissement de la GSF (semaine 23 début juin), comme le montre la figure 2.

En ce qui concerne la faune particulière de la mi-septembre, nous l'avons rapprochée de l'homogénéisation verticale de la couche d'eau; ceci peut se vérifier par l'examen des densités en surface et au fond. L'isolement, essentiellement sur l'axe 2 des semaines 36, et surtout 37 et 38 est spécialement net pour les années 1969, 1971 et 1973.

On peut vérifier (Merle, Guillermin, 1975) que ce sont précisément celles où l'homogénéisation verticale est la plus spectaculaire. On remarquera même l'exacte simultanéité des deux phénomènes.

### Année moyenne en quantitatif

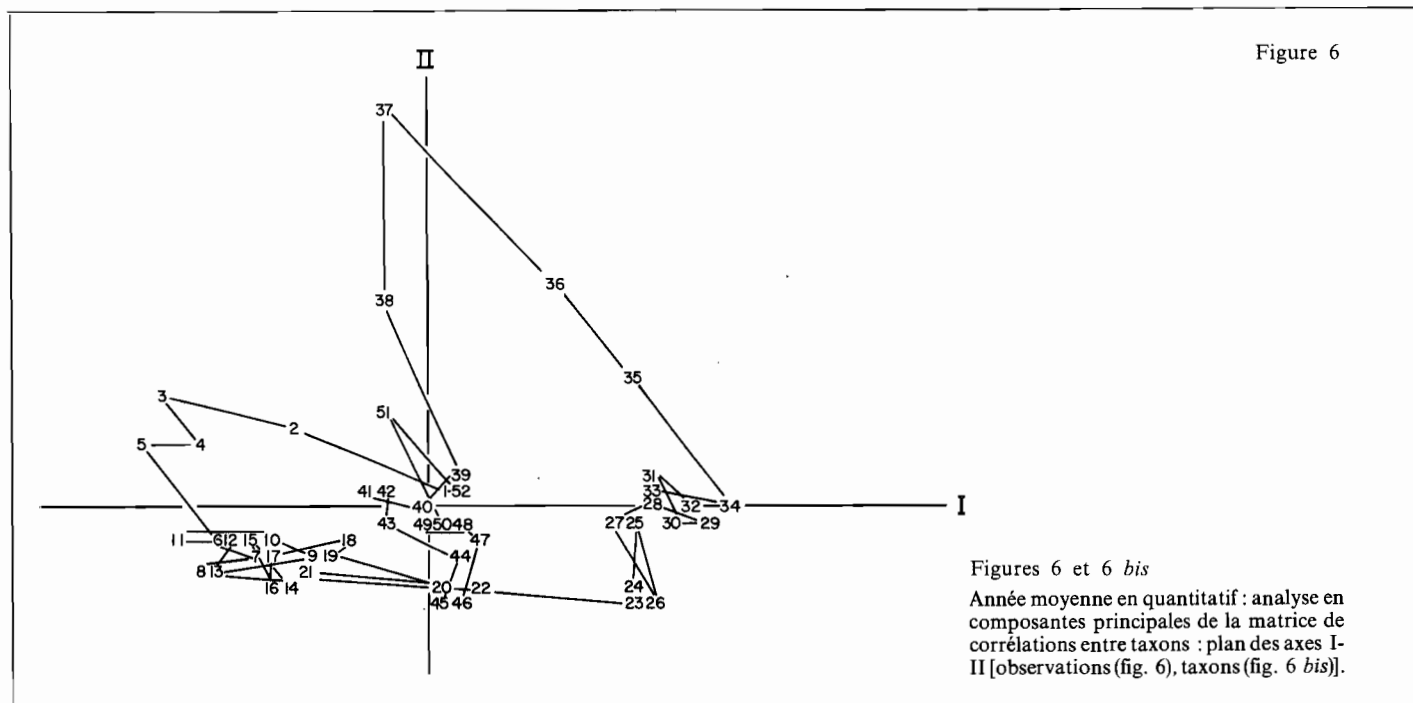
#### Corrélation entre espèces

Comme annoncé précédemment une analyse en composantes principales a été pratiquée sur la matrice des corrélations entre espèces, issue du tableau représentant l'année moyenne.

#### • Cycle saisonnier

Les traits généraux du cycle tel qu'ils ont été mis en évidence dans l'étude semi-quantitative sont conservés. On retrouve les mêmes phénomènes et les premiers axes des deux études sont associés aux mêmes structures. Les pourcentages d'inertie sont quelque peu différents, étant cette fois respectivement pour les quatre premiers axes : 29, 12, 9 et 5 %. Nous observons donc essentiellement une opposition entre les grandes saisons, chaudes et froides, tandis que les petites saisons sont relativement peu individualisées à l'encontre des époques de transition (associées au début ou à la fin des grandes saisons). Cette convergence est d'autant plus significative qu'elle reste vraie si l'on se limite à une étude proprement qualitative

Figure 6



en présence-absence. La figure 6 montre bien la concordance des résultats. Toutefois le début juin (la semaine 23 notamment), correspondant à l'installation de la GSF, est beaucoup moins bien individualisé. L'étude semi-quantitative avait montré que cette individualisation était liée essentiellement à des présences, plutôt qu'à des modifications d'abondance. Il est donc normal qu'une telle structure soit estompée dans une étude quantitative. En revanche, l'isolement sur le deuxième axe des semaines voisines de la 37<sup>e</sup>, relié à la transition de la GSF vers la PSC, est nettement plus marqué (fig. 6). Les petites saisons apparaissent également plus individualisées dans le plan des axes III et IV qui permet de mieux les étudier (fig. 7).

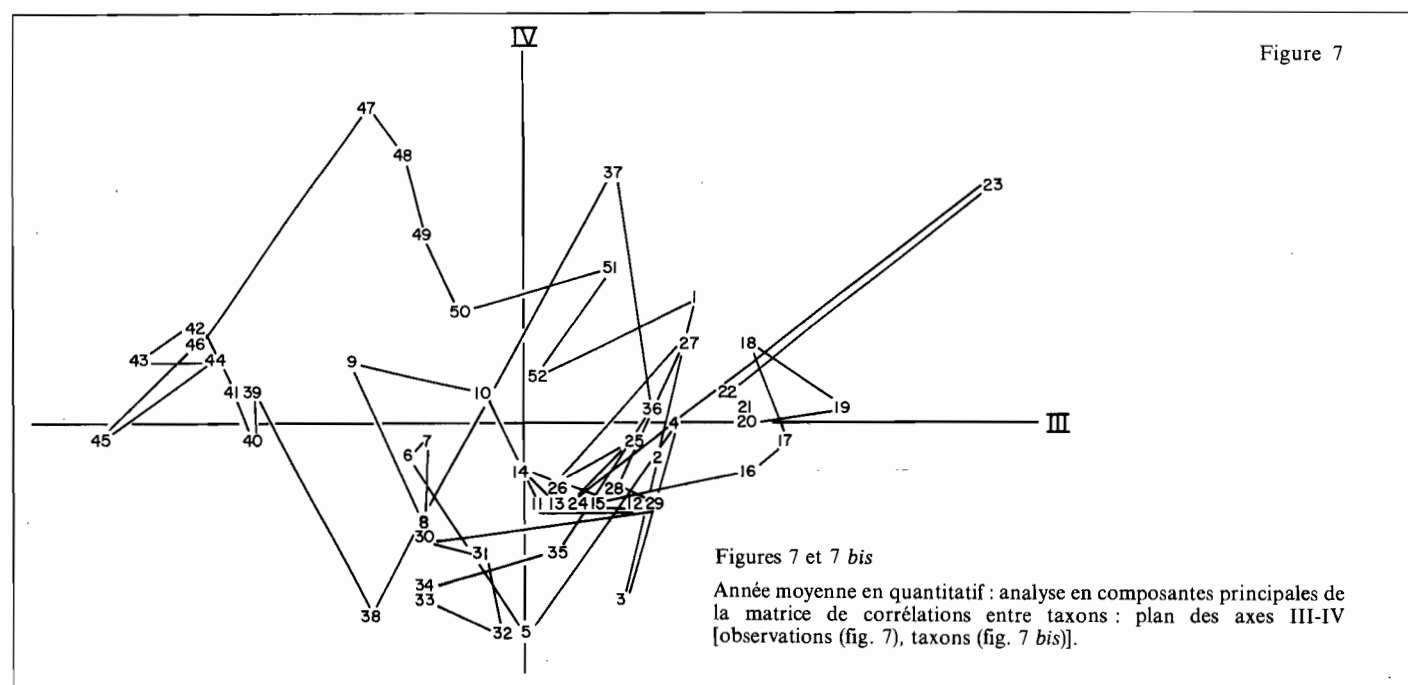
• Étude de l'espace dual des taxons

Là encore, il n'apparaît pas de grande nouveauté.

L'interprétation est dans l'ensemble moins riche, comme on l'a annoncé, puisqu'il n'existe plus qu'une variable par taxon. Cependant, les taxons ayant été réduits, les masses étant désormais égales pour toutes les variables celles-ci ont toutes la même contribution totale. Il n'y a plus d'ambiguïté au niveau de l'interprétation de l'éloignement d'une variable par rapport à un axe comme significatif de son rôle prééminent dans la détermination de cet axe.

On peut comparer les figures 2 bis et 6 bis, d'abord au niveau des taxons dont l'étude semi-quantitative avait pris en compte les fluctuations d'abondances (essentiellement ceux ayant donné naissance à trois états et représentés par des cercles sur la figure 2 bis). La concordance est alors flagrante, tant au niveau de ceux dont l'abondance culmine en GSC (*Paracalanus*

Figure 7





dans l'analyse semi-quantitative. Elles se détachent sur la figure 7 assez nettement dans le quatrième quadrant. Une certaine cassure apparaît entre la 46<sup>e</sup> et la 47<sup>e</sup> semaine qui correspond (fig. 6) à un rebroussement sur l'axe 2. La figure 7 bis montre que celle-ci est essentiellement liée à des taxons méroplanctoniques ou, de façon plus générale, à des œufs et larves (principalement *E. pileatus*, nauplii de copépodes — appartenant dans une très grande majorité à l'espèce *E. pileatus* — larves de décapodes, Zoés de brachyours, œufs et larves de poissons et, plus spécialement pour la PSF, les larves actinotroques et celles de stomatopodes). Quelques autres taxons font preuve d'une certaine préférence pour les petites saisons (tel *E. pileatus* pour la PSF). Là encore l'étude semi-quantitative manquait un peu de nuances, car les taxons qui viennent d'être évoqués n'avaient été étudiés qu'en présence-absence (ou plus généralement sur deux états). En réalité un examen plus approfondi montre que les taxons qui apparaissent liés aux petites saisons sur la figure 6 bis n'ont pas nécessairement leur abondance maximale à cette époque. Ce sont en général des taxons tout aussi abondants en GSC, mais ne reflétant pas la situation intermédiaire des petites saisons, ainsi que cela peut s'observer sur la figure 6.

#### Covariance entre espèces

L'analyse en composante principale pratiquée sur la matrice de covariance et non des corrélations entre espèces, nous amène à restituer une influence prépondérante aux taxons numériquement prédominants. On effectue alors, comme nous l'avons signalé, un premier pas vers une approche plus fonctionnelle. L'inconvénient majeur reste peut-être à cet égard l'influence excessive accordée aux taxons de la base de la pyramide alimentaire, qu'il conviendrait de pondérer. Ceci, sans même envisager les difficultés qu'il y a à situer un taxon dans les chaînes alimentaires, nous entraînerait trop loin. De plus et compte tenu des conditions très particulières de notre station, l'importance numérique d'un taxon ne présume pas forcément une situation en début de chaîne alimentaire. Le fait intéressant réside en ce que la modification de point de vue (attachée à l'utilisation de la distance euclidienne avec des variables non réduites) n'entraîne pas de modification sensible dans les grandes lignes de la structure du cycle saisonnier. Ceci suggère que les facteurs essentiels au terme des études précédentes (telle que la température), sont aussi ceux qui gouvernent la variation des structures trophiques des peuplements. Cette concordance n'est pas une trivialité, et Chardy *et al.* (1976) ont montré comment les résultats des analyses pouvaient varier selon le point de vue adopté.

En ce qui concerne la structure duale des taxons, seuls ceux numériquement abondants présentent une contribution significative. Cette structure duale montre que la GSF est alors caractérisée par un nombre très réduit de taxons qui atteignent des abondances considérables, essentiellement *Paracalanus scotti* et *Oithona nana*, à un degré moindre *Pseudodiaptomus serricaudatus* et *Acartia plumosa*. En revanche, beaucoup plus nombreux sont les taxons caractérisant la GSC. L'isolement des semaines 36-37-38 et 2-3-4 est principalement dû à

*Penilia* qui présente pour la 37<sup>e</sup> semaine un effectif moyen de 4 100 individus par prélèvement, alors qu'il n'atteint pas le dixième de cette valeur en GSF ou en PSC. Un accroissement similaire, mais moins prononcé peut être observé en janvier. *Temora turbinata* apparaît aussi, à un degré moindre comme prépondérant dans la transition PSF-GSC.

## CONCLUSIONS

Les conclusions qui peuvent être formulées à l'issue d'une telle étude sont de trois sortes. On peut tout d'abord tenter la synthèse, d'un point de vue écologique, des résultats obtenus. On peut ensuite évoquer les points essentiels de la méthodologie de l'analyse mathématique des données et enfin poser le problème de la stratégie de l'échantillonnage sur le terrain, à suivre dans d'éventuelles études similaires futures.

### Écologie

D'un point de vue général, nous retrouvons les grandes lignes des saisons marines telles qu'elles avaient été décrites par Berrit (1958). L'apport essentiel de l'étude présente, en ce qui concerne le cycle du zooplancton, est la mise en évidence de la très forte singularité, au moins sur le plan écologique, des périodes de transition de PSF à GSF (janvier) et surtout de GSF à PSC (septembre). Ces phénomènes n'ont cependant pu être reliés à des événements hydrologiques évidents. Leur interprétation nécessiterait vraisemblablement une étude d'ensemble du plateau continental, tout comme la dissymétrie constatée entre les transitions de saison chaude à saison froide et de saison froide à saison chaude. Une première conséquence de cette analyse est l'apport possible de l'étude des populations zooplanctoniques dans une description des phénomènes hydrologiques saisonniers. L'abondance (voire simplement la présence) d'une espèce intègre les effets d'un très grand nombre de paramètres de son environnement (non seulement présent mais aussi passé), nombre dépassant très largement celui généralement mesuré. Le problème posé par l'utilisation des abondances d'espèces zooplanctoniques se situe au niveau de leur estimation *in situ*.

Outre l'intérêt présenté par le zooplancton en tant que révélateur de phénomènes hydrologiques (établissement des conditions saisonnières, intensité des saisons marines, etc.) et que nous avons abondamment illustré, cette étude a montré la grande simplicité de structure du cycle moyen. Les premiers axes intègrent l'essentiel de l'information et nous avons vu la parfaite liaison de l'axe I avec la température, paramètre du milieu le plus accessible et reflétant le mieux les phénomènes hydrologiques. On a pu remarquer la faible individualisation des petites saisons; ce résultat est moins immédiat qu'il y paraît; le passage des peuplements de GSC à ceux de GSF se fait par remplacement progressif sans qu'apparaisse un nombre important de taxons qui, à l'exclusion des situations saisonnières extrêmes, préfèrent les conditions moyennes des petites saisons. Ainsi le facteur général, lié à la température, joue-t-il un rôle essentiel quoique d'une grande simplicité.



La liaison constatée entre le cycle zooplanctonique moyen et les phénomènes hydrologiques peut être étendue aux « anomalies » apparaissant dans le cycle saisonnier. Nous avons montré comment les « accidents » sur l'axe I pouvaient être immédiatement rapprochés de variations thermiques. En fait, il ne faut pas perdre de vue la situation très particulière de la station, quelque peu en marge des écosystèmes pélagiques se succédant au cours de l'année (en même temps que les différentes masses d'eau) au-dessus du plateau continental congolais et dont l'étude nécessitera l'examen des prélèvements recueillis à une vaste échelle géographique pendant les campagnes du NO « André Nizery ». Toute perturbation du milieu se traduit non seulement par une modification des paramètres hydrologiques (en premier lieu la température et liées à celle-ci la teneur en phosphates dissous, la salinité, etc.) mais aussi par l'arrivée de populations exogènes venant se substituer, ou plus généralement s'ajouter aux populations pré-existantes. La quasi-simultanéité de ces perturbations de l'environnement physicochimique et de leur manifestation sur l'axe I ne peut évidemment s'expliquer par une adaptation instantanée du peuplement zooplanctonique à de nouvelles conditions. De nombreux exemples peuvent venir illustrer ces assertions : un refroidissement modéré intervenant à la faveur d'un petit upwelling temporaire, fréquent en cours de GSC, se traduira aussi par l'augmentation concomitante des effectifs de certaines espèces (telles *Paracalanus scotti* ou *Oithona nana*, dont l'abondance caractérise la GSF) manifestement transportées par les eaux issues de l'upwelling; un réchauffement généralisé de toute la couche pourra par contre être attribué à l'arrivée jusqu'à la côte d'eaux guinéennes chaudes et peu salées et de la faune qui y est associée. L'auto-écologie des espèces et taxons fait mieux apparaître les particularités de la station qu'une étude synécologique, et invite à beaucoup de prudence dans l'interprétation des phénomènes généraux tels qu'ils ressortent de l'analyse. Le retour à une situation « normale » (pour la saison) après qu'aient cessé de se manifester les effets d'une perturbation du milieu traduit davantage le retour des peuplements chassés ou modifiés par des apports exogènes, que la robustesse d'un peuplement permanent à l'égard des modifications de son environnement. De même, la réapparition saisonnière de peuplements identiques résulte-t-elle principalement du passage à Pointe-Noire des mêmes masses d'eau auxquelles ils sont inféodés.

Il n'en reste pas moins vrai que chaque année aux mêmes conditions hydrologiques correspondent des peuplements comparables. Un exemple particulièrement frappant en est fourni par les périodes de transition surtout celles de GSF à PSC (en septembre) où l'on voit les mêmes taxons caractériser très fortement ces périodes.

On pourrait craindre à ce niveau un artéfact, l'année moyenne ayant conduit à créer une période caractérisée par un certain nombre de taxons, alors qu'au niveau des années individuelles il y aurait vicariance des espèces caractéristiques, donc non reproductibilité. Le retour aux données de base, comme l'analyse globale des 601 prélèvements montrent que la reproductibilité est réelle.

On ignore à peu près tout de l'histoire des masses d'eau (et des peuplements qu'elles abritent) qui passent devant Pointe-Noire à la transition GSF-PSC. On peut raisonnablement supposer que cette histoire diffère d'une année sur l'autre. Un écart, même très faible, pourrait entraîner une divergence dans l'évolution ultérieure des peuplements. Une telle divergence, si elle existe, est probablement d'amplitude et de durée limitées puisque l'examen ponctuel, que permet le passage au wharf de ces eaux montre une reproductibilité remarquable. Il n'est hélas pas possible de préciser l'importance et le mécanisme des phénomènes de régulation qui permettent d'absorber au niveau du zooplancton de très probables accidents, assurant la stabilité des équilibres écologiques.

## Méthodologie du traitement

### Codage

Le codage logique permet essentiellement de respecter la singularité de l'information apportée par une présence ou une absence, sans renoncer totalement à celle liée aux variations d'abondance. Au niveau de l'interprétation, ceci apporte une finesse certaine, en marquant ce qui caractérise un groupe de prélèvements. Si l'étude était centrée sur les relations entre variables avant de l'être sur la structure d'un ensemble de prélèvements, le codage logique présenté serait une excellente façon de traiter le problème de relations non monotones, dont l'importance a été soulignée notamment par van Groenewoud (1976).

Cette approche comporte cependant des inconvénients. La définition des variables utilise des histogrammes empiriques. Les résultats des analyses ne sont donc pas directement comparables d'un jeu de données à l'autre. La présentation des résultats sous une forme standard peut pourtant être, comme le notent Blanc et Laurec (1976) un des avantages essentiels du traitement mathématique des données. Par ailleurs, comme on l'a déjà souligné, le codage logique entraîne une perte d'information. Ceci justifie le recours parallèle à une approche quantitative. Il est vain de chercher l'approche uniformément optimale. La combinaison d'études passant par un codage logique, avec d'autres utilisant des variables quantitatives constitue en fait la meilleure démarche.

### Apports des analyses factorielles

L'utilité de ces techniques a pu être mise en doute (Frontier, 1975); ceci est justifié si l'on se réfère à l'approche classique. L'approche suivie dans cet article se veut essentiellement descriptive, comme le suggèrent Blanc et Laurec (1976). Si l'on veut se souvenir des buts assignés à l'étude, et des espoirs placés au début de cet article, en les analyses d'inertie, on pourra constater que ces espoirs n'ont pas été déçus. Les analyses ont bien fourni des synthèses caractérisant les structures essentielles du cycle saisonnier, structures qui ont pu être reliées aux taxons. On peut même remarquer qu'un fait non antérieurement signalé a été révélé par l'analyse : l'originalité de la transition GSF-PSC. La région fut dans le passé relativement peu étudiée. Une analyse classique, non mathématique des données, aurait probablement conduit à relever le phénomène évoqué. Il est impossible

de savoir dans quel délai, compte tenu de la masse des données à dépouiller. Sur des données bien connues, les analyses mathématiques recoupent généralement les connaissances antérieures (Frontier, 1975). Sur des sujets moins déflorés, on voit qu'elles permettent de faire très rapidement une synthèse et offrent par là un gain de temps.

Le regroupement de prélèvements avant l'analyse permet de focaliser l'étude sur des sources de variations choisies. Le procédé comporte toutefois des dangers que l'on devra garder à l'esprit. La combinaison avec la technique des points témoins permet, dans l'étude des phénomènes saisonniers, une forme intéressante d'étude des anomalies. Hors même de ce contexte, il nous paraît que l'utilisation d'éléments supplémentaires constitue une technique qui mériterait d'être plus souvent utilisée en écologie marine qu'elle ne l'a été jusqu'à présent.

Si les analyses factorielles nous apparaissent comme des techniques précieuses, rien ne serait plus dangereux que limiter les traitements mathématiques à l'utilisation de ces seuls outils. A cet égard, l'annexe C constitue une réflexion sur l'articulation de l'analyse présentée avec d'autres études.

### Stratégie d'échantillonnage

La définition d'une stratégie d'échantillonnage doit être rapportée aux problèmes d'inférences, tels qu'ils sont esquissés dans l'annexe B. Dans cette annexe deux thèmes sont évoqués : le nombre d'années d'études, et l'influence couplée des variations à très courts termes et des phénomènes de microdistribution.

Deux autres aspects sont à étudier : le nombre de taxons à prendre en compte, la précision dans l'estimation de leur abondance.

Nous supposons au long de la discussion que les taxons sont utilisés essentiellement comme révélateurs. Nous nous en tiendrons aux faits majeurs. Il serait dangereux de leur donner une valeur universelle, compte tenu des particularités de la station. Des études analogues, actuellement en cours, et portant sur des sites côtiers de mers tempérées, conduisent cependant dans l'ensemble aux mêmes conclusions.

La première question a trait au nombre des taxons à considérer. Il n'est pas possible au niveau d'analyse où nous nous situons d'apprécier parfaitement l'information que peuvent apporter les taxons les plus rares. Elle dépend pour une bonne part des particularités du site étudié. Ainsi malgré une situation très néritique de notre station, la configuration de la côte, ainsi que la largeur relativement modeste du plateau continental, l'ouvrent aux influences hauturières. Ces influences peuvent se concrétiser selon la saison par l'arrivée à la côte d'espèces subsuperficielles lors de l'établissement des upwellings, ou d'espèces superficielles des eaux chaudes du large en période estivale. Il demeure que la prise en compte d'un nombre élevé de taxons reste un élément de succès. Ce sont généralement des complexes de facteurs liés qui sont envisagés et l'étude simultanée d'un ensemble d'espèces pouvant être sensibles à l'une ou l'autre des variables élémentaires permettra de mieux dégager ces facteurs

complexes. Il existe d'autre part une certaine indépendance dans les fluctuations d'échantillonnage correspondant aux divers taxons, ce qui aura, sur un plan statistique, un effet stabilisateur. Ceci joue surtout s'il y a homogénéité du rôle des variables. Dans une analyse en composante principale sur une matrice de variance-covariance, pour déborder sur des analyses plus fonctionnelles, seuls les taxons abondants définissent en pratique les structures. La stabilisation des résultats par le nombre des taxons étudiés est beaucoup moins sensible.

Le second problème est relatif à la précision requise dans l'estimation de l'abondance d'un taxon, les études en présence-absence constituant un cas extrême. Nous avons vu que les structures déduites d'analyses semi-quantitatives étaient, pour l'essentiel celles obtenues par une analyse quantitative. Cette constatation appelle deux remarques. Tout d'abord, une étude n'est jamais totalement qualitative, surtout lorsque sont créés des tableaux de contingence; si une espèce est présente dans le milieu, mais rare, elle pourra apparaître ou non dans l'échantillon, puis dans le sous-échantillon, sa fréquence d'apparition étant alors liée à sa densité dans le milieu.

D'autre part, une étude en présence-absence peut manquer de nuance car si une espèce est soumise à un facteur quelconque, hormis un certain intervalle de transition où elle pourra être ou non présente, elle sera systématiquement présente ou absente. Les valeurs critiques pouvant différer d'une espèce à l'autre, les nuances seront restituées dans une étude multispécifique mais il demeure évident que des finesses de structure seront perdues si l'on s'en tient seulement aux présences-absences. L'information quantitative requise pour les déceler est malheureusement très coûteuse; un moyen terme pourra être alors le recours à une cotation rapide d'abondance (Frontier, 1969). Il nous paraît toutefois judicieux de coupler une telle cotation avec l'utilisation de codages logiques, plutôt que d'opérer un traitement quantitatif comme le fait par exemple Ibanez (1974).

Le troisième problème concerne la répartition dans le temps des échantillons; deux aspects sont à considérer : la maille temporelle et la durée de l'étude. Les populations de zooplancton évoluent de façon relativement lente malgré les différences sensibles entre les régions tempérées et les régions intertropicales. Il est peu probable que des changements significatifs apparaissent d'un jour à l'autre. A une station fixe ce qu'on observe essentiellement c'est le défilé et le mélange de masses d'eau. A ce niveau des changements rapides peuvent intervenir mais d'une année sur l'autre ces changements ne se reproduisent pas à date fixe. En cherchant à dégager une année moyenne, on opère un lissage qui rend vain l'étude de structures temporelles très fines. Un échantillonnage hebdomadaire nous semble raisonnable dans l'exemple que nous proposons. Utiliser une maille plus fine peut se justifier, d'un point de vue statistique, afin de réduire les fluctuations d'échantillonnage, et d'accroître la fiabilité de l'estimation des lignes du tableau d'abord, des structures finales ensuite.

En ce qui concerne la durée de l'étude, même si la reproductibilité d'une année sur l'autre est dans notre exemple relativement frappante, des anomalies ont pu

être signalées. Il serait imprudent qu'elle fût brève, une seule année est certainement insuffisante. Si un choix devait être fait, nous suggérerions plutôt d'accroître la maille temporelle, ou de diminuer la précision de l'analyse des récoltes, jusqu'à une étude en présence-absence, plutôt que de limiter l'étude à une seule année. Bien entendu, une telle proposition pourrait être critiquable en d'autres lieux.

## RÉFÉRENCES

- Benzecri J. P. et al.**, 1973. *L'analyse des données*, II, *L'analyse des correspondances*, Dunod, Paris, 619 p.
- Berrit G. R.**, 1958. Les saisons marines à Pointe-Noire, *Bull. Coec (Cah. Océanogr. et Études des Côtes)*, **10**, 6, 336-358.
- Binet D., Dessier A., Gaborit M., Roux M.**, 1972. Premières données sur les copépodes pélagiques de la région congolaise, *Cah. Orstom, sér. Océanogr.*, **10**, 2, 125-137.
- Blanc F., Chardy P., Laurec A., Reys J. P.**, 1976. Choix des métriques qualitatives en analyse d'inertie. Implications en écologie marine benthique, *Mar. Biol.*, **35**, 49-67.
- Blanc F., Laurec A.**, 1976. De l'heuristique au thaumaturgique en traitement des données d'écologie marine, *Cah. Orstom, sér. Océanogr.*, **XIV**, 2, 101-107.
- Cailliez F., Pages J. P.**, 1976. *Introduction à l'analyse des données*, Smash, 616 p.
- Cassie R. M.**, 1963. Multivariate analysis in the interpretation of numerical plankton data, *N.Z.J. Sci.*, **2**, 239-365.
- Chardy P., Glemarec M., Laurec A.**, 1976. Application of Inertia Methods to Benthic Marine Ecology: practical implications of the basic options, *Estuarine Coastal Mar. Sci.*, **4**, 170-205.
- Cordier B.**, 1965. *L'analyse des correspondances*. Thèse, Doc. Sci., Rennes, 100 p.
- Dufour P., Merle J.**, 1972. Station côtière en Atlantique tropical. Hydroclimat et production primaire, Orstom Doc. Sci. du centre de Pointe-Noire, N.S. n° 25, 48 p., 21 fig.
- Fasham M. J. R.**, 1977. A comparison of non-metric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes, *Ecology*, **58**, 3, 551-561.
- Frontier S.**, 1969. Sur une méthode d'analyse faunistique rapide du zooplancton, *J. Exp. Mar. Biol. Ecol.*, **3**, 18-26.
- Frontier S.**, 1973. Étude statistique de la dispersion du zooplancton, *J. Exp. Mar. Biol. Ecol.*, **12**, 3, 229-262.
- Frontier S.**, 1975. L'analyse factorielle est-elle heuristique en écologie du plancton? *Cah. Orstom, sér. Océanogr.*, **XII**, 1, 77-81.
- Gauch H. G., Wittaker R. H., Wentworth T. R.**, 1977. A comparative study of reciprocal averaging and other ordination techniques, *J. Ecol.*, **65**, 157-174.
- Gower J. C.**, 1966. The distance properties of latent roots and vector methods used in multivariate analysis, *Biometrika*, **53**, 325-338.
- Guillerm J. M., Le Bouteiller A.**, 1972. Observations océanographiques côtières Pointe-Noire. Année 1970, Orstom Doc. Sci. Centre Pointe-Noire, N.S. n° 21, 50 p., 11 fig.
- Hirshfeld H. O.**, 1935. A connection between correlation and contingency, *Proc. Cambridge Philos. Soc.*, **31**, 520-524.
- Hotelling H.**, 1933. Analysis of a complex of statistical variables into principal components, *J. Educ. Psychol.*, **29**, 417-441 et 498-520.
- Ibanez F.**, 1974. Une cotation rapide d'abondance réduite à trois classes : justification de son emploi en analyse des composantes principales. Mise en œuvre et intérêt pratique en planctonologie, *Ann. Inst. Océanogr. Paris*, **50**, 2, 185-198.
- Ibanez F., Seguin G.**, 1972. Étude du cycle annuel du zooplancton d'Abidjan. Comparaison de plusieurs méthodes d'analyse multivariées : composantes principales, correspondances, coordonnées principales, *Inv. Pesq.*, **36**, 1, 81-108.
- Lebart L., Morineau A., Tabard N.**, 1977. *Techniques de la description statistique*, Dunod, Paris, 352 p.
- Merle J., Guillerm J. M.**, 1975. Recueil provisoire de graphiques d'évolution annuelle à la station du wharf (Pointe-Noire) 1969-1974, document interne, antenne Orstom-Cob.
- Pielou E. C.**, 1969. *An introduction to mathematical ecology*, Wiley Interscience, New York, 286 p.
- Prentice I. C.**, 1977. Non-metric ordination methods in ecology, *J. Ecol.*, **65**, 85-94.
- Reys J. P.**, 1976. Les peuplements benthiques (zoobenthos) de la région de Marseille (France) : aspects méthodologiques de la délimitation des peuplements par les méthodes mathématiques, *Mar. Biol.*, **36**, 123-134.
- Reyssac J., Roux M.**, 1972. Communautés phytoplanctoniques dans les eaux de la Côte-d'Ivoire. Groupe d'espèces associées, *Mar. Biol.*, **13**, 14-34.
- Van Groenewoud H.**, 1976. Theoretical considerations on the covariation of plant species along ecological gradients with regard to multivariate analysis, *J. Ecol.*, **64**, 837-847, 7 fig.
- Verneaux J.**, 1973. Cours d'eau de Franche-Comté (Massif du Jura). Recherches écologiques sur le réseau hydrographique du Doubs. Essai de biotypologie, *Ann. Sci. Univ. Besançon (Zool.)*, **9**, 1-257, 61 fig.

## ANNEXE A

### CHOIX D'UNE ANALYSE FACTORIELLE PARTICULIÈRE

#### Critères de choix

Le souci d'analyse fine des liens entre la structure de l'ensemble des semaines et les taxons conduit à se limiter aux structures euclidiennes simples, où l'on ne joue que sur des transformations préliminaires éventuelles du tableau analysé, et sur une pondération des lignes et colonnes.

A l'intérieur même de cette catégorie d'analyse des choix doivent être effectués qui pourront notamment conduire aux variantes les plus classiques (analyse des correspondances, analyse en composantes principales).

Plusieurs études ont été consacrées à la comparaison de l'utilisation en écologie de ces techniques particulières (Chardy *et al.*, 1976; Gauch *et al.*, 1977). Le point essentiel n'est alors pas de rechercher une méthode qui serait universellement optimale, mais de comprendre ce qu'apporte chaque variante, pour choisir la technique la

mieux adaptée à un problème. Dans cet esprit, l'accent étant mis sur la structure de l'ensemble des semaines, le point le plus important correspond au choix de la distance entre semaines. Cette distance est définie pour un tableau donné par des transformations préliminaires et une pondération des variables, mais le choix du procédé de codage intervient également : c'est toute une filière, incluant le codage, qui doit être analysée. Ceci doit être fait sur chaque cas particulier, car les mêmes formules mathématiques peuvent selon les cas avoir une signification concrète différente.

Une distance ayant été choisie, pour pratiquer une analyse d'inertie, les différentes observations doivent recevoir des masses. Ceci constitue le second point à analyser. Un troisième aspect est dans l'absolu envisageable (Chardy *et al.*, 1976), celui du point à partir duquel est pratiqué l'analyse d'inertie. Dans les études présentées, pour bénéficier de la propriété de représentation optimale, ce sera toujours le barycentre du nuage des observations.

Lorsque l'on en viendra au second propos des analyses d'inertie, l'étude du lien des variables avec les structures de l'ensemble des semaines, il n'y a pas d'ambiguïté au

niveau de l'interprétation des contributions. Il pourra y en avoir lorsque l'on tentera de percevoir les configurations variables comme des représentations de ces contributions. Au niveau des contributions des variables aux axes cette utilisation ne sera pas possible lorsque les variables auront des masses inégales. On pourra encore réfléchir dans chaque cas particulier sur l'interprétation des configurations variables comme traduisant des liens préférentiels entre certaines variables et certains groupes d'observations. Enfin, on devra s'interroger sur la possibilité d'utiliser les mêmes configurations variables comme rendant compte des proximités entre variables. Cette réflexion portera sur deux points : d'une part on examinera la distance entre variables impliquée par l'analyse (c'est-à-dire par la transformation préalable du tableau et les masses attribuées aux observations), d'autre part on examinera si les configurations variables constituent ou non des représentations optimales.

### Étude semi-quantitative et analyse des correspondances

Le tableau constitué est directement apparenté à une table de contingence. L'analyse des correspondances, Hirschfeld (1935), Fisher (1940), Cordier (1965), et Benzecri (Benzecri *et al.*, 1973) est un outil spécialement adapté. Soit le tableau principal, excluant les éléments supplémentaires.  $x_{i.}$  et  $x_{.j}$  correspondent respectivement à la somme de la ligne  $i$  et de la colonne  $j$ , et définiront la masse attribuée à ces éléments. Une transformation préliminaire du tableau sera effectuée  $x_{i,j}$  étant dans un premier temps divisé par le produit  $x_{i.} \cdot x_{.j}$ . Le centrage des lignes, qui amène l'origine au barycentre des semaines, vient ultérieurement compléter cette première transformation, sans affecter les distances entre semaines. La distance entre les semaines  $i_1$  et  $i_2$  est donnée par la distance dite du  $\chi^2$  :

$$d^2_{i_1, i_2} = \sum_j \frac{1}{x_{.j}} \left( \frac{x_{i_1, j} \cdot x_{i_2, j}}{x_{i_1.} \cdot x_{i_2.}} \right)^2 \\ = \sum_j x_{.j} \left( \frac{x_{i_1, j}}{x_{i_1.} \cdot x_{.j}} - \frac{x_{i_2, j}}{x_{i_2.} \cdot x_{.j}} \right)^2.$$

Elle correspond à la combinaison des transformations préalables évoquées avec l'attribution du poids  $x_{.j}$  à la variable  $j$  (Chardy *et al.*, 1976).

Le codage pratiqué étant disjonctif complet, au sens que, pour tout prélèvement, parmi les conditions attachées à un taxon, une et une seule est toujours satisfaite,  $x_{i.}$  ne dépend pas de la semaine  $i$ . La division par  $x_{i.}$  est donc sans impact. Au niveau de la distance entre semaines, le coefficient  $1/x_{.j}$  accroît l'influence des événements rares, comme l'apparition d'un taxon peu fréquent.

Par ailleurs, la distance du  $\chi^2$  possède la propriété dite d'équivalence distributionnelle (Benzecri *et al.*, 1973). Deux conditions logiques sont équivalentes si elles engendrent des colonnes proportionnelles. On peut alors les réunir pour définir une condition globale remplie dès que l'une des deux est satisfaite. Ceci amène à sommer les deux colonnes. La propriété d'équivalence assure que les distances entre semaines ne seront pas affectées. Si pour

un taxon le découpage de la plage de variation fut inutilement fin, parce que créant deux colonnes proportionnelles, ceci n'affecta donc pas les distances entre semaines.

L'analyse des correspondances suppose l'attribution à la ligne  $i$  de la masse  $x_{i.}$ . Les  $x_{i.}$  étant toutes égales, ceci est sans impact particulier. L'éloignement d'une variable sur un axe n'implique pas une contribution importante, car sa masse  $x_{.j}$  peut être faible. On devra donc consulter le tableau des contributions des variables aux axes.

L'interprétation simultanée des configurations des semaines et des variables dans des plans analogues sera rendue spécialement fructueuse par la propriété dite barycentrique : moyennant un choix approprié d'échelles, dans une projection simultanée des observations et des variables, la variable  $j$  apparaîtra au barycentre des diverses semaines, avec un poids  $x_{i,j}/x_{i.}$  pour la semaine  $i$ . Seules notamment les semaines où  $x_{i,j}$  est non nulle interviendront. Les variables reliées à une période précise seront clairement marquées.

La distance entre variables est également donnée par la distance du  $\chi^2$ .

$$d^2_{j_1, j_2} = \sum_i \frac{1}{x_{i.}} \left( \frac{x_{i, j_1}}{x_{.j_1}} - \frac{x_{i, j_2}}{x_{.j_2}} \right)^2.$$

La division de chaque  $x_{i,j}$  par  $x_{.j}$  correspond à un passage de fréquences absolues à des fréquences relatives. Le coefficient  $1/x_{i.}$  est par ailleurs sans impact. Le centrage des colonnes équivaut en analyse des correspondances à celui des lignes. Les configurations des points variables constitueront donc des représentations optimales.

En marge de l'analyse du tableau de base, les années individuelles ont été reportées comme éléments supplémentaires. Les distances entre semaines, moyennes ou individuelles, sont données par la distance du  $\chi^2$ , mais les lignes supplémentaires n'interviennent pas dans la définition de  $x_{.j}$ , somme de la colonne  $j$  du tableau de base.

L'analyse des correspondances a été utilisée en écologie marine par de nombreux auteurs, Reyssac et Roux (1972), Binet *et al.* (1972), Ibanez et Segin (1972), Chardy *et al.* (1976). La présente étude se singularise par la combinaison avec un codage logique, ce qui constitue un domaine privilégié d'application de l'analyse des correspondances. On remarquera à ce propos que l'utilisation d'un codage disjonctif complet amène des particularités intéressantes, détaillées par Lebart *et al.* (1977) à propos de l'analyse des correspondances multiples.

### Études quantitatives et analyses en composantes principales

On fait traditionnellement remonter ces techniques à Hotelling (1933). Nous avons cependant suivi une approche descriptive qui n'est pas celle des premiers travaux, même si l'on pratique les mêmes calculs mathématiques.

*Analyse de la matrice de corrélation des taxons* : les variables, c'est-à-dire les colonnes, sont préalablement centrées et réduites. Il n'y a pas de pondération ultérieure, lignes et colonnes ont toutes des masses égales.

La distance entre deux semaines est donnée par

$$d^2_{i_1, i_2} = \sum_j \left( \frac{x_{i_1, j}}{sc_j} - \frac{x_{i_2, j}}{sc_j} \right)^2,$$

$sc_j$  étant l'écart-type de la colonne  $j$ .

Comme les variables ont des masses égales, les configurations variables indiquent directement les contributions aux axes. Pour interpréter simultanément les configurations duales on ne disposera pas d'un résultat aussi simple que la propriété barycentrique. En revanche, on pourra interpréter les configurations variables en utilisant le fait que la coordonnée d'une variable par rapport à un facteur n'est autre que sa corrélation avec ce facteur. On peut ainsi marquer qu'un axe, qui isole un groupe de semaines, est ou non corrélé avec l'abondance d'un taxon.

La distance entre deux variables  $j_1$  et  $j_2$  est donnée par

$$d^2_{j_1, j_2} = 2(1 - co_{j_1, j_2}), \quad co_{j_1, j_2}$$

désignant leur coefficient de corrélation.

Cette distance peut être considérée comme une mesure intéressante du lien entre variables. En revanche, les configurations variables ne constituent pas des représentations optimales.

*Analyse de la matrice de covariance-covariance des taxons* : la seule différence avec l'analyse précédente concerne le fait que les variables sont centrées mais non réduites. Sur le plan écologique, ce point est important : ceci aboutira à accorder aux taxons abondants une influence majeure (Pielou, 1969). Ce peut être justifié si l'on tend vers une étude plus fonctionnelle des peuplements : il n'est pas déraisonnable de considérer que les taxons numériquement abondants puissent jouer un rôle prépondérant dans l'économie de l'écosystème. Dans une telle approche on ne dispose cependant que d'une information très insuffisante, puisque limitée à une estimation de l'effectif des taxons. Dans l'idéal, il faudrait connaître leur production, et utiliser une pondération adaptée. Il faudrait encore disposer des abondances réelles, qui seraient utilisées sans transformation le problème statistique ne se poserait plus, et la transformation de l'échelle est sans justification dans une approche fonctionnelle. En ce qui concerne ce second point, la transformation  $\text{Log}^2(x+1)$  présente le mérite d'être relativement peu écrasante.

Ce n'est qu'un tout premier pas qui peut être effectué vers une approche fonctionnelle. Il n'est pourtant pas sans intérêt de vérifier si l'évolution du point de vue entraîne ou non une distorsion considérable des résultats.

Au niveau des particularités de l'interprétation, ce qui fut dit de l'analyse de la matrice des corrélations reste vrai, à ceci près que les coordonnées des variables ne seront plus des coefficients de corrélation, mais de simples covariances.

Les analyses en composante principale sont d'un usage relativement ancien en écologie marine (Cassie, 1963), sans que l'accent ait toujours été mis sur l'aspect descriptif. Si au lieu de suivre l'approche descriptive adoptée ici, on tente d'expliquer une matrice de corrélation ou de covariance selon l'approche historiquement antérieure, des difficultés apparaissent très vite en écologie marine (Frontier, 1975). Il paraît donc préférable de suivre l'approche descriptive (Blanc, Laurec, 1976).

## ANNEXE B

### NOTES SUR LES PROBLÈMES D'INFÉRENCE STATISTIQUES

L'utilisation des analyses factorielles comme des techniques de description d'un tableau bidimensionnel, repousse au second plan les problèmes d'inférence statistique. Il est évident cependant que l'analyse présentée ne vise pas seulement à décrire les structures de l'ensemble des prélèvements étudiés, mais bien à étudier le cycle saisonnier du zooplancton, tel qu'il se déroule au wharf de Pointe-Noire. Il existe donc une inférence, mais le problème se pose dans des termes différents des analyses factorielles classiques.

#### Définition du doublet échantillon population

Si dans les tentatives d'inférence l'échantillon est le plus souvent clairement défini, car correspondant à l'ensemble des prélèvements étudiés, il n'en va pas toujours de même de la population à laquelle s'étend l'inférence (Blanc, Laurec, 1976). Les analyses factorielles classiques mettent l'accent sur les relations entre variables à l'intérieur d'un bloc spatio-temporel, constituant la population, les prélèvements correspondant à autant de réalisations indépendantes de la loi multivariable sous-jacente. Ce point de vue est inadapté dans la plupart des problèmes d'écologie marine. Loin d'être indépendants les prélèvements correspondent à une structure qui fait l'objet de l'étude. Dans le problème étudié on doit considérer que, les variables ayant été déterminées, on souhaiterait connaître, pour toute semaine d'un cycle annuel moyen, la valeur moyenne pour cette semaine de chaque variable. Ceci constituerait le tableau que l'on souhaiterait en fait pouvoir analyser. On ne dispose que d'estimations des lignes du tableau. Ces estimations sont d'abord entachées par la variabilité des résultats d'un prélèvement à l'intérieur d'une semaine, correspondant aux variations à très courts termes, et surtout aux problèmes de microdistribution. Disposer de plusieurs prélèvements, ultérieurement moyennés, réduit le problème, mais ne le supprime pas. Sur les estimations jouent encore les variations d'année en année. Dans cet esprit, parler d'année moyenne, sans la référer à une période précise, suppose implicitement que les phénomènes sont stationnaires, sans tendance au fil des années. Ceci pourrait être discuté.



### Estimation des distances entre semaines

En utilisant pour estimer les distances entre semaines la distance entre les estimations des lignes on sera nécessairement conduit à des estimations biaisées comme on le vérifie aisément. Si le biais est sensiblement le même pour tous les couples, ceci amènera un gonflement global des distances mutuelles, sans modifier l'écart entre elles. L'inégalité des biais est beaucoup plus gênante. Les études quantitatives sont extrêmement sensibles à cet aspect. Comme l'a exposé le texte principal, ceci conduit à utiliser des transformations préalables, qui auront pour but d'éviter que les distances entre semaines riches soient considérablement plus surestimées que celles entre semaines pauvres.

### Impact global sur le résultat des analyses

Puisque les lignes du tableau ne constituent que des estimations, le résultat des analyses, c'est-à-dire les configurations des semaines et des variables dans les premiers axes, et les inerties correspondantes, ne constituent eux-aussi que des estimations. On entre ainsi dans une catégorie de problèmes analogues à ceux étudiés par Tarek Moussa (*in Benzecri et al.*, 1973). Pour fonder une inférence sur les résultats des analyses il faudrait connaître la loi multivariable des erreurs d'estimations du tableau. Peu d'études furent consacrées à ce problème; en ce qui concerne le phénomène essentiel de microdistributions, les recherches sont généralement centrées sur l'aspect univariable. A défaut d'une impossible analyse exacte, nous avons examiné la stabilité des résultats lorsque l'on diminuait soit le nombre d'années étudiées, soit le nombre de prélèvements dans chaque semaine. Les faits commentés dans l'article sont conservés, et se retrouvent même lorsque l'on procède à l'analyse sur deux sous-échantillons disjoints.

Un autre problème se pose encore, celui de la détermination du nombre d'axes retenus à l'issue de l'analyse. La limitation est venue de nos capacités d'interprétation des configurations, non de considérations statistiques. Les axes suivant les tous premiers ne paraissent pas correspondre à un phénomène particulier. Ils interviennent pour corriger et nuancer les phénomènes liés aux tous premiers axes. Au vu des inerties leur impact est de plus très limité.

## ANNEXE C

### ARTICULATION AVEC D'AUTRES ÉTUDES

Le présent article n'épuise pas l'information contenue dans les données recueillies au long du programme Wharf. D'autres paramètres ont été collectés que la température, et les variables zooplanctoniques. Par ailleurs une optique descriptive a délibérément été choisie; l'aspect dynamique n'a pas été développé. Les

techniques d'analyse d'inertie ne sont guère adaptées à une telle approche. Dans les premiers axes, il serait concevable de placer l'interprétation sous un signe dynamique, en étudiant les traits essentiels des évolutions. Il demeure que les analyses utilisées ne prennent pas en compte la structure particulière, temporelle, de l'ensemble des prélèvements. Si l'on désire étudier la dynamique fine des peuplements, on doit faire appel à d'autres techniques, et notamment à l'analyse des séries temporelles. Dans cet esprit l'analyse présentée peut constituer une phase préliminaire : les techniques classiques supposent les séries étudiées sans tendance. Lorsque des tendances existent elles doivent être soustraites et étudiées séparément. Le cycle saisonnier entraîne l'existence d'une telle tendance, que l'on peut considérer comme périodique. En déterminant une année moyenne, on a estimé cette tendance, et l'étude présentée est une description, sous un angle multivariable, des caractéristiques essentielles de la tendance. Une telle étude doit être complétée par l'analyse des écarts vis-à-vis de la tendance. L'étude de telles anomalies a été esquissée par l'utilisation des points témoins. Ceci ne peut dégager que les faits essentiels. Un traitement par les techniques d'analyses liées aux séries temporelles serait en théorie mieux à même de préciser les aspects dynamiques. Une telle étude a été entreprise, qui intègre des variables hydrologiques et hydrobiologiques. Elle sera présentée dans un article ultérieur. On remarquera toutefois que les techniques d'analyse des séries temporelles sont rapidement coûteuses, lorsque l'on s'adresse à un nombre élevé de variables. L'analyse qui vient d'être présentée peut contribuer à dégager un nombre réduit de variables synthétiques, telles celle définie par l'axe 1, qui fut utilisée sur la figure 5. Là encore il n'y a pas compétition entre les techniques d'étude, mais complémentarité.

Sans détailler les résultats de l'analyse des séries temporelles, on peut rapidement constater qu'au niveau du zooplancton on ne doit pas en attendre de résultats fondamentaux sur la dynamique des peuplements. Comme on l'a déjà souligné ce que l'on suit au wharf c'est essentiellement un défilé de masse d'eaux, non une évolution interne des peuplements. Si par des techniques d'analyse des séries temporelles on étudie les relations entre le facteur I et la température, on peut confirmer qu'un lien existe entre les anomalies, même pour les phénomènes de haute fréquence, et qu'aucun délai significatif n'apparaît.

Il y a confirmation utile d'une impression visuelle. En revanche, les hypothèses de stationnarité font que l'on perd la possibilité d'une étude spécifique de chaque anomalie précise : on étudie des phénomènes moyens. Au mieux, pourra-t-on étudier séparément les différentes saisons. Par ailleurs, il est difficile d'étudier simultanément dans une série temporelle deux variables, sans utiliser des modèles linéaires non nécessairement adéquats. Sur ce point, pour étudier les anomalies combinées des facteurs I et II il est difficile de se passer d'une étude graphique, telle qu'elle fut effectuée par les points témoins.