

TRAITEMENT DES DONNEES FOURNIES PAR L'ECHANTILLONNAGE
DU MACROBENTHOS DES SUBSTRATS MEUBLES.

par

Raphaël PLANTE

Station Marine d'Endoume, 13007 Marseille

(France)

SUMMARY

As an attempt to summarize the main types of methods for classification and ordination of benthic samples, the author gives a critical analysis of criteria for the choice of specific and cenotic indexes and of methods for the use of matrices.

For a valuable exploitation of data, several methods should be used at the same time. If data are not normally distributed, convenient methods must be chosen. Complexity of original data should often be reduced by changing quantitative results into qualitative ones.

INTRODUCTION

Dans le cas le plus général, après collecte et tri des échantillons de macrobenthos, on dispose d'un ou plusieurs tableaux donnant la représentation (qualitative, quantitative ou pondérale) des différentes espèces dans les différentes stations. Le but du présent article est de commenter les méthodes qui permettent d'utiliser ces matrices de données brutes en vue de la description des associations benthiques. En se cantonnant à la description (c'est-à-dire en excluant des études de production, par exemple), on peut, en effet, rechercher:

- la définition des ensembles biocénotiques à partir de données d'ordre qualitatif ou plutôt d'ordre qualitatif relatif.
- la définition des rapports structuraux existant entre différents groupes de stations (dans certains cas, cette catégorie rejoint la précédente) à partir de données quantitatives.
- la définition du degré d'évolution de différents écosystèmes, sans faire intervenir la structure des peuplements proprement dite mais plutôt la diversité spécifique.

Chacune de ces démarches a reçu de nombreux développements. La première (étude qualitative des biocénoses) a été largement utilisée sur les rives méditerranéennes selon le processus défini par PICARD

(1965). Nous nous intéresserons plutôt aux méthodes mettant en jeu des outils mathématiques. On en trouvera une revue complète et critique dans BLANC et REYS (en préparation) et nous nous contenterons ici de quelques remarques sur leurs limites et sur l'intérêt qu'ils présentent.

1 . CHOIX DES INDICES DE LIAISON CENOTIQUES ET SPECIFIQUES.

Quelles que soient les techniques d'analyse dont on envisage de se servir (ce choix étant souvent conditionné par l'accès à un ordinateur), l'un des points les plus importants est certainement le choix de la *métrique* c'est-à-dire de l'indice de liaison qui servira de système de mesure.

La différenciation des groupes de stations⁽¹⁾ peut, en effet, se faire à l'aide d'indices dont la multiplicité traduit bien la complexité des critères selon lesquels une telle classification peut être faite.

Sur le plan de la nature des données de départ, on reconnaît, parmi ces coefficients et indices, trois catégories essentielles (BLANC et al. 1976):

- 1- Les indices qualitatifs, fondés sur des critères de présence ou absence.
- 2- Les indices quantitatifs, fondés sur des critères de fréquence, d'abondance ou de dominance.

-3- Les indices semi-quantitatifs, fondés sur le rang des espèces.

De même, si on les classe selon la signification des modèles mathématiques, on peut distinguer:

- a- Les distances ou indices de distance (les plus connues sont sans doute la distance euclidienne et la distance du *Chi carré*).

(1) ... ou des groupes d'espèces: en effet, une série d'observations peut être interprétée au niveau de la liaison entre les espèces dans l'ensemble des stations, ou, complémentairement, de la liaison entre les stations sur l'ensemble des espèces (notion de dualité).

-b- Les coefficients de dépendance: les divers coefficients de corrélation et de régression.

-c- Les indices de similitude, comme, par exemple, les indices de JACCARD et de SORENSEN.

Il est impossible de recommander un indice d'usage universel; le choix devra se faire en fonction des critères énoncés ci-dessus et en tenant compte des points suivants:

1 . 1 . Exigences des indices

Les indices de similitude sont souvent utilisés en raison de leur signification écologique bien nette. Ils font intervenir, par exemple, selon des modalités très diverses, les nombres d'espèces présentes dans les deux prélèvements, absentes dans les deux, ou présentes dans l'un des deux. Elaborés dans un but déterminé, ils utilisent de façon particulière l'information de départ. Ce caractère empirique restreint leur utilisation.

C'est pourquoi on leur préfère souvent des coefficients de corrélation comme le coefficient de BRAVAIS-PEARSON. Mais l'utilisation de celui-ci exige que les données de départ soient des données quantitatives distribuées selon la loi de LAPLACE-GAUSS. De même, l'usage des corrélations partielles et multiples suppose une multinormalité des données. Ces conditions sont rarement satisfaites dans les peuplements naturels, de sorte qu'on devra remédier à ce défaut par deux types de solutions:

- ou bien on "normalisera" les données de départ par des transformations appropriées (les plus employées sont de la forme $\log x$, n^x).

- ou bien on utilisera une métrique (un indice) qui ne présuppose pas une distribution normale, c'est-à-dire faisant appel à des statistiques non paramétriques: coefficients de corrélation de rang, coefficient de KRUSKAL-WALLIS.

1 . 2 . Codage des données

Dans le cas, le plus courant, où l'on cherche à déterminer le rôle des différentes espèces dans le regroupement des stations, les in-

dices privilégient les espèces abondantes, ou bien la présence des espèces. Or, l'absence dans une station d'une espèce abondante ailleurs, peut avoir une signification que l'on pourra mettre en évidence par un codage adéquat des données. Dans ce cas particulier, on établira une "matrice-miroir" des données: le code de départ étant 1 pour la présence et 0 pour l'absence, le code utilisé dans la matrice-miroir sera 1 pour l'absence et 0 pour la présence.

Ces transformations et codages de données sont une étape importante dans l'analyse. On peut être tenté, pour conserver toute l'information contenue dans les données de départ, de les conserver sous leur forme quantitative primitive, mais l'on s'expose alors au danger d'appliquer des modèles inadéquats (cf. § 1.1). En fait, l'utilisation d'une partie seulement de l'information de départ permet de mieux mettre en évidence les phénomènes fondamentaux en supprimant les "bruits de fond" (REYS, 1976; DESSIER et LAUREC, 1978). Une telle simplification peut s'obtenir par plusieurs méthodes:

- transformation des données quantitatives en données qualitatives. Une solution intéressante consiste par exemple à utiliser un code à 5 niveaux, comme l'emploient les phytosociologues:

0: absent, 1: rare, 2: peu abondant,
3: abondant, 4: très abondant.

- regroupement d'espèces ou de stations
- suppression des espèces peu représentées
- utilisation de métriques simplificatrices

2 . EXPLOITATION DES MATRICES D'INDICES

Les matrices obtenues peuvent être utilisées de façon plus ou moins élaborée selon l'importance des moyens de calcul dont on dispose. Une des techniques les plus sommaires, très largement utilisée, consiste, à partir de matrices carrées (symétriques) dont on manipule les colonnes de façon convenable intuitivement, à mettre en évidence des groupements d'espèces ou de stations (*diagrammes-treillis*). Cependant, ces méthodes deviennent très difficiles à employer au-dessus d'un cer-

tain nombre de stations et d'espèces.

Si l'on dispose d'un ordinateur, les méthodes d'analyse deviennent presque aussi nombreuses que les indices (Tableau 1) et il ne peut être question de les exposer toutes ici.

Tableau 1. - Principales méthodes d'analyses utilisées pour l'étude des peuplements benthiques.

METHODES D'ORDINATION	METHODES DE CLASSIFICATION
1-TECHNIQUES FACTORIELLES	1-METHODES DIVISIVES
- En composantes principales	- Formation de classes homothétiques
- Factorielles <i>sensu stricto</i>	- Segmentation
- Rotations orthogonales ou obliques	- Formation de classes polythétiques
- Analyse des correspondances	- Méthode de FAGER (Recurrent groups)
- En coordonnées principales	2-METHODES AGGLOMERATIVES
- Analyse factorielle des matrices de corrélations résiduelles après partialisation multiple	a-Méthodes hiérarchiques
2-ANALYSE DISCRIMINANTE (Normalité des données)	-M. du plus proche voisin
3-CORRELATION CANONIQUE	-M. du voisin le plus éloigné
	-M. de la liaison moyenne
	-M. de la liaison moyenne pondérée
	-Flexible strategy
	-Analyse de l'information
	-Somme des carrés
	b-Méthodes non hiérarchiques
	-M. des constellations (distance D^2 de MAHALANOBIS)
	-M. des nuées dynamiques

Nous renvoyons le lecteur à BLANC et REYS (en préparation). Le choix

final dépend de l'objectif poursuivi: les méthodes de classification conviendront s'il s'agit de mettre en évidence des groupes distincts (biocénoses, communautés), alors que les méthodes d'ordination seront mieux adaptées pour définir la place d'une observation dans une série continue (continuum) (BOUDOURESQUE, 1970).

C'est là, évidemment, une description simpliste, mais il est cependant vrai que les méthodes de classification sont tout à fait adéquates dans la description des associations en vue d'une cartographie, alors que les techniques d'ordination et d'analyse d'inertie permettent d'éclairer le rôle de tel facteur ou de telle espèce.

Les méthodes d'analyse factorielle, par exemple, peuvent montrer la contribution des différentes espèces à l'explication du phénomène global et là encore, il peut être intéressant de manipuler les données de départ pour aller plus loin dans l'analyse: il arrive qu'à un premier degré de l'analyse on mette en évidence une espèce ou un facteur qui joue un rôle prépondérant, et dont le "poids" écrase celui des autres variables; si tel est le cas, on aura avantage à continuer l'analyse sur la matrice débarrassée de cette variable, afin de mieux préciser les détails secondaires de la structure.

Au stade actuel, il semble bien que les techniques factorielles soient les plus utilisées par les benthologues (MONBET, 1972; GUILLE et PONGE, 1975; REYS, 1976), si l'on excepte les travaux de l'école australienne, davantage fondés sur des méthodes de classification hiérarchique (STEPHENSON, 1973; STEPHENSON et al., 1972; BOESCH, 1973).

3 . LA DIVERSITE SPECIFIQUE ET L'EQUILIBRE DU MILIEU

La façon dont, dans un biotope, les individus sont répartis entre les espèces traduit à l'évidence, un degré d'organisation plus ou moins élaboré de l'association étudiée. Pour quantifier cette complexité de l'écosystème, plusieurs techniques de mesure de la diversité spécifique existent. Des revues critiques qu'en on fait TRAVERS (1971) et MARGALEF (1977), on peut conclure que les plus fréquemment utilisées sont:

- des indices mesurant l'uniformité ou la non-uniformité de

distribution (indices de SIMPSON ou de MORISITA).

- des indices fournissant une évaluation de la "néguentropie" d'un système; cette évaluation du "désordre" est très séduisante et les formules les plus connues sont celles de MARGALEF et de SHANNON-WEAVER. Une telle mesure est souvent perfectionnée par l'évaluation de l'"équitabilité" (contribution des abondances relatives des différentes espèces).
- des techniques graphiques (SANDERS, 1968; FRONTIER, 1969) extrêmement intéressantes quand on cherche (c'est le cas le plus fréquent) à comparer entre eux plusieurs milieux ou prélèvements homologues.

Un des espoirs qui expliquent le succès de cette notion de diversité spécifique repose sur l'hypothèse d'une relation entre cette *diversité* et la *stabilité* de l'écosystème (SANDERS, 1969). En fait, la justesse de cette hypothèse est maintenant contestée (BOESCH, 1972; GOODMAN, 1975), mais beaucoup d'auteurs s'en servent comme indice d'équilibre ou de déséquilibre d'un milieu, notion d'un intérêt pratique évident dans les études de pollution par exemple.

4 . DISCUSSION - CONCLUSIONS

= Bien que ce ne soit pas le propos exact de cet article, il nous faut rappeler que les résultats d'une analyse sont fortement conditionnés par la nature des données de départ, et que le plan d'échantillonnage a donc une importance extrême en benthologie, la nature de cet échantillonnage pouvant introduire des artefacts au niveau de l'analyse (IBÁÑEZ, 1965).

= Etant donnée la multiplicité des situations que l'on peut avoir à étudier, il n'est pas possible de recommander un ou des indices particuliers. Il faut se souvenir que chacun d'entre eux filtre l'information de départ de façon particulière.

= Par conséquent, il est nécessaire d'appliquer simultanément, et de façon critique, plusieurs méthodes d'analyse au même faisceau de données.

= Les indices fondés sur des critères statistiques, de même que certaines

méthodes d'analyse, exigent une normalité des données de base. Si cette normalité n'existe pas, il peut être bon de recourir à des indices tels que les corrélations de rang, et à des techniques d'analyse moins contraignantes à l'égard de la normalité comme l'Analyse des Correspondances.

= Dans le but de clarifier les images obtenues, il est parfois souhaitable de simplifier la structure de départ en rendant par exemple qualitatives, toutes les données quantitatives, même s'il en résulte la perte d'une certaine quantité d'information.

REFERENCES BIBLIOGRAPHIQUES

- BLANC(F.),CHARDY(P.),LAUREC(A.) et REYS (J.P.),1976.-Choix des métriques qualitatives en analyse d'inertie. Implications en écologie marine benthique.*Mar. Biol.*35 (1) pp. 49-67.
- BLANC(F.) et REYS(J.P.).- Outils mathématiques et délimitation des communautés. *in* Marine Ecology . O.Kinne ed. (à paraître).
- BOESCH (D.F.), 1972.-Species diversity of marine macrobenthos in the Virginia area. *Chesapeake Sci.* 13, pp.206-211.
- BOESCH (D.F.),1973.- Classification and Community structure of macrobenthos in the Hampton Roads area, Virginia. *Mar. Biol.*21(2)pp. 226-244.
- BOUDOURESQUE(C.F.),1970.-Recherches sur les concepts de biocoenose et de continuum au niveau des peuplements benthiques sciaphiles. *Vie et Milieu, sér.B*,21 (1) pp.103-136.
- DESSIER(A.)et LAUREC(A.),1978.- Le cycle annuel du zooplancton à Pointe-Noire (R.P.Congo). Description mathématique.*Acta Oceanol.*,1 (3) pp.285-304.
- FRONTIER(S.),1969.- Méthodes d'analyse statistiques applicables à l'écologie du plancton.*Doc. Sci. Centre ORSTOM Nosy-Bé*,7, 33p.
- GOODMAN (D.),1975.- The theory of diversity-stability relationships in ecology. *Quart. Rev. Biol.*,50 (3) pp.237-266.
- GUILLE (A.) et PONGE (J.F.),1975.- Application de l'analyse des correspondances à l'étude des peuplements benthiques de la côte catalane française.*Ann. Inst. Océanog.* 51 (2) pp. 223-230.
- IBANÉZ (F.),1969.- Application de l'analyse factorielle en planctologie: écologie et taxinomie numérique. Thèse 3ème cycle Univ. Paris. Multigr. 130p.
- MARGALEF (R.),1977.- Ecologia. 2ème ed. 951p. Barcelona.
- MONBET (Y.),1972.- Etude bionomique du plateau continental au large d'Arcachon (Application de l'analyse factorielle). Thèse 3ème cycle. Univ. Aix-Marseille II. Multigr. 98p.
- PICARD (J.),1965.- Recherches qualitatives sur les biocénoses marines des substrats meubles dragables de la région marseillaise. *Rec. Trav. St. mar. Endoume.* 52 (Bull. 36. 160 p.).
- 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 (2). pp. 123-134.
- SANDERS (H.L.), 1968.- Marine benthic diversity: a comparative study. *Amer. Natural.* 102, pp. 243-282.
- 1969.- Benthic marine diversity and the stability-time hypothesis. *Brookhaven Symp. Biol.* 22 pp. 71-81.
- STEPHENSON (W.), 1973.- The validity of the community concept in marine biology. *Proc. R. Soc. Qd.* 84 (7) pp. 73-86.
- STEPHENSON (W.), WILLIAMS (W.T.) et COOK (S.D.), 1972.- Computer analysis of Petersen's original data on bottom communities. *Ecol. Monogr.* 42, pp. 387-415.
- TRAVERS (M.), 1971.- Diversité du microplancton du Golfe de Marseille en 1964. *Mar. Biol.* 8 (4). pp. 308-343.

