
Inv. Pesq.	36 (1)	págs. 109-118	junio 1972
------------	--------	---------------	------------

Analyse en composantes principales de données de phytoplancton de la zone côtière du Sud de l'Ebre

par

MARTA ESTRADA *

Dans ce travail on expose quelques résultats de l'application de méthodes d'analyse multivariable à une série de données de phytoplancton de l'aire côtière du Sud de l'Ebre.

Les échantillons ont été recoltés pendant une traversée de Barcelona à Castelló, à bord du bateau nord-Américain Thompson. Les prises et les comptages des échantillons ont été réalisés par le Professeur R. Margalef. On a disposé aussi de mesures de chlorophylle (pour la moitié des stations), de fluorescence et de quelques paramètres chimiques : silicates, nitrates, phosphates. Les ordinateurs utilisés ont été le IBM 360 du Service de Calcul de l'Université de Barcelona et le IBM 1130 de la Faculté des Sciences.

Les positions de prise des échantillons sont marquées sur les cartes (fig. 4-7) ; la profondeur était toujours d'environ 1 m.

On a négligé la technique d'échantillonnage au hasard étant donné les possibilités du voyage en mer. La disposition en zig-zag a été choisie parce qu'on désirait cartographier la distribution des paramètres mesurés, et le temps de navigation était limité. Le nombre total de prélèvements obtenus a été 67.

Les nombres de cellules recomptées ont été référés à 100 ml ; il faut signaler que pour certaines espèces le nombre d'individus comptés était très bas, quelquefois un seulement.

Des 87 espèces trouvées au total on en a choisi 42 pour une analyse en composantes principales ; on a limité le nombre à cause de la capacité

* Ecología, Facultad de Ciencias, Universidad de Barcelona.

de l'ordinateur et du temps de calcul. On a éliminé les espèces qui avaient été signalées uniquement dans un ou deux échantillons et avec une abondance très basse (table I).

On n'a pas fait des tests de normalité pour la distribution de chacune des espèces ; car la transformation logarithmique ($X \rightarrow \log(X+1)$) réussit souvent à normaliser des données d'abondance de ce type et possède une certaine justification théorique, on l'a donc appliqué dans ce cas.

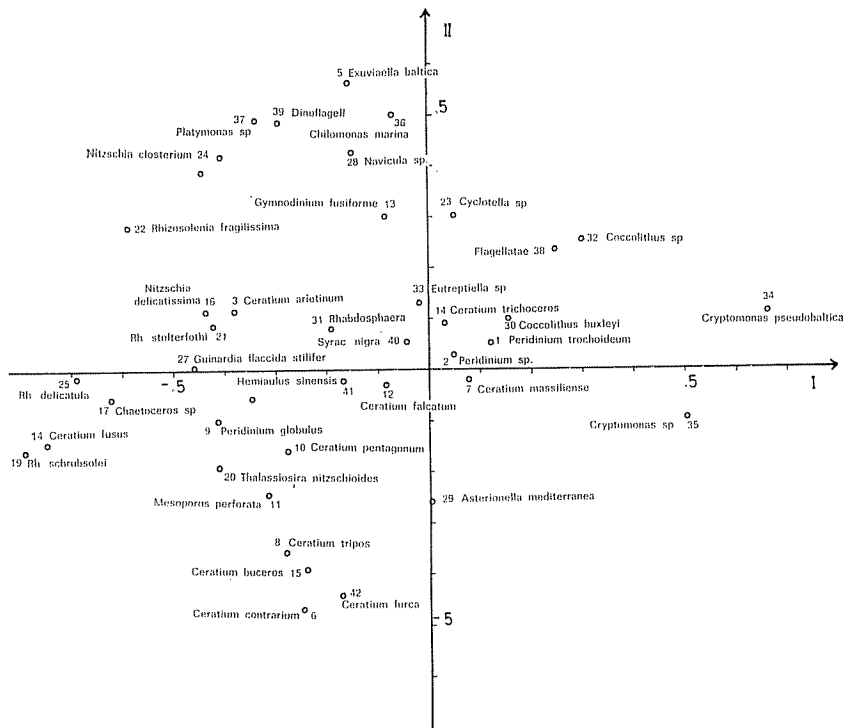


FIG. 1. — Position des espèces dans l'espace des axes I-II. Les coordonnées des points sont les «factor loadings» correspondants.

Dans une première analyse on a compté parmi les 42 espèces deux groupes : «petits dinoflagellés» et «petits flagellés», avec respectivement des petits dinoflagellés et flagellés qui n'avaient pu être identifiés. On a fait une deuxième analyse en éliminant ces deux groupes et substituant quelques espèces par d'autres qui n'avaient pas été considérées la première fois. Les résultats étaient très semblables ; on ne considérera ici que ceux de la première analyse.

L'analyse en composantes principales a été faite sur la matrice de corrélation ; on n'a appliqué aucune rotation. Le nombre de valeurs pro-

pres obtenues supérieures à l'unité a été 16 ; les pourcentages accumulés de variance expliquée par les composantes correspondantes sont :

13,59	20,74	26,98	32,65	38,03	42,56	47,01	51,10
54,93	58,59	62,13	65,44	68,38	71,29	73,96	76,44

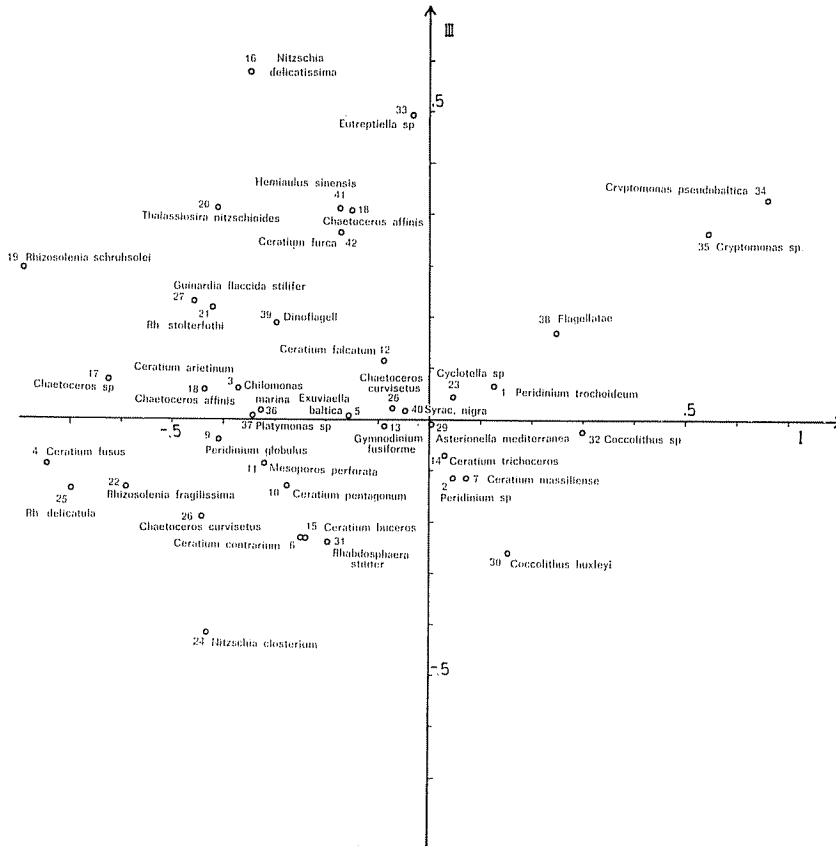


FIG. 2. — Position des espèces dans l'espace des axes I-III. Les coordonnées des points sont les «factor loadings» correspondants.

Ce qui frappe déjà ici est la petite partie de la variance expliquée par les premières composantes. En ce sens un résultat comparable est celui d'une analyse en composantes principales de phytoplancton des Caraïbes (MARGALEF et GONZÁLEZ BERNÁLDEZ, 1969), où l'on a trouvé que les 4 premières composantes expliquaient 12,85 % de la variance.

On peut comparer aussi les résultats obtenus avec ceux de l'étude par BLASCO (1971) de 18 espèces du phytoplancton des côtes du Peru.

T A B L E I

Liste des espèces et coefficients de charge pour les 4 premières composantes

	1	2	3	4
1. Peridinium trochoideum	0,123	0,051	0,066	0,128
2. Peridinium sp.	0,052	0,034	0,155	0,106
3. Ceratium arietinum	-0,378	0,113	0,060	-0,524
4. Ceratium fusus	-0,748	-0,151	-0,089	0,162
5. Exuviaella baltica	-0,156	0,561	0,004	0,155
6. Ceratium contrarium	-0,247	-0,457	-0,223	0,634
7. Ceratium massiliense	0,779	-0,020	-0,118	0,224
8. Ceratium tripos	-0,330	-0,360	0,016	-0,138
9. Peridinium globulus	-0,413	-0,103	-0,036	-0,315
10. Ceratium pentagonum	-0,276	-0,166	-0,130	-0,427
11. Mesoporus perforata	-0,317	-0,240	-0,084	0,156
12. Ceratium falcatum	-0,086	-0,039	0,118	0,096
13. Gymnodinium fusiforme	-0,085	0,304	-0,011	0,093
14. Ceratium trichoceros	0,033	0,092	-0,074	0,048
15. Ceratium buceros	-0,241	-0,395	-0,230	0,592
16. Nitzschia delicatissima	-0,346	-0,058	0,680	0,157
17. Chaetoceros sp.	-0,626	-0,057	0,081	-0,092
18. Chaetoceros affinis	-0,436	0,115	0,059	-0,009
19. Rhizosolenia schrubsolei	-0,792	-0,169	0,296	0,013
20. Thalassiosira nitzschioides	-0,411	-0,243	0,414	0,001
21. Rhizosolenia stolterfothi	-0,421	0,087	0,229	-0,155
22. Rhizosolenia fragilissima	-0,592	0,279	-0,135	0,309
23. Cyclotella sp.	0,052	0,306	0,046	-0,047
24. Nitzschia closterium	-0,433	0,415	-0,427	-0,033
25. Rhizosolenia delicatula	-0,698	0,022	-0,138	-0,006
26. Chaetoceros curvisetus	-0,440	0,390	-0,192	-0,179
27. Guinardia flaccida	-0,460	0,003	0,230	0,338
28. Navicula sp.	-0,149	0,420	0,413	0,314
29. Asterionella mediterranea	0,007	-0,258	-0,011	-0,005
30. Coccolithus huxleyi	0,157	0,101	-0,227	-0,002
31. Rhabdosphaera stilifer	-0,194	0,080	-0,239	0,067
32. Coccolithus sp.	0,300	0,257	-0,242	-0,012
33. Eutreptiella sp.	-0,028	0,132	0,597	-0,237
34. Cryptomonas pseudobaltica	0,666	0,113	0,436	0,244
35. Cryptomonas sp.	0,546	-0,097	0,367	0,132
36. Chilomonas marina	-0,073	0,503	0,018	0,124
37. Platymonas sp.	-0,339	0,485	0,011	0,099
38. Flagellatae	0,253	0,237	0,170	0,092
39. Dinoflagellatae	-0,297	0,483	0,189	-0,029
40. Syracosphaera nigra	-0,044	0,053	0,018	0,311
41. Hemiaulus sinensis	-0,169	-0,026	0,414	0,251
42. Ceratium furca	-0,171	-0,440	0,365	-0,363

Dans cette étude, on a trouvé que les cinq premières composantes expliquaient 71,1 % de la variance, et on a pu interpréter d'une façon biologiquement très raisonnable les groupements d'espèces trouvés; il faut signaler que, dans ce cas, on avait choisi, parmi les 103 espèces déterminées, ces 18 espèces, en essayant d'avoir des représentants de tous les groupes et avec une abondance importante. En outre, étant donné

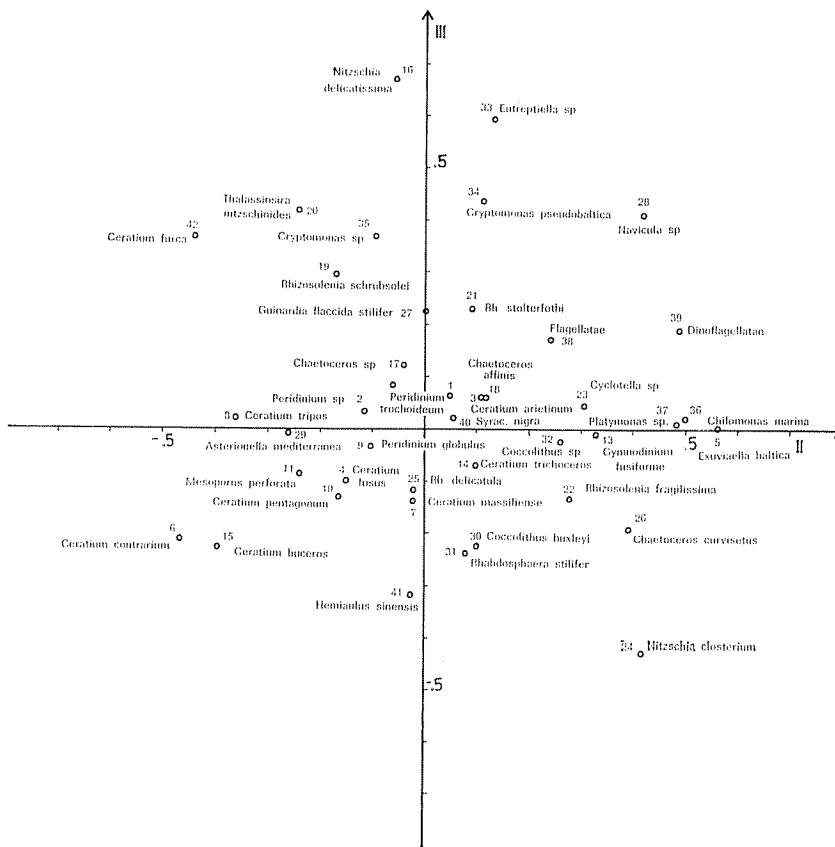


FIG. 3. — Position des espèces dans l'espace des axes II-III. Les coordonnées des points sont les «factor loadings» correspondants.

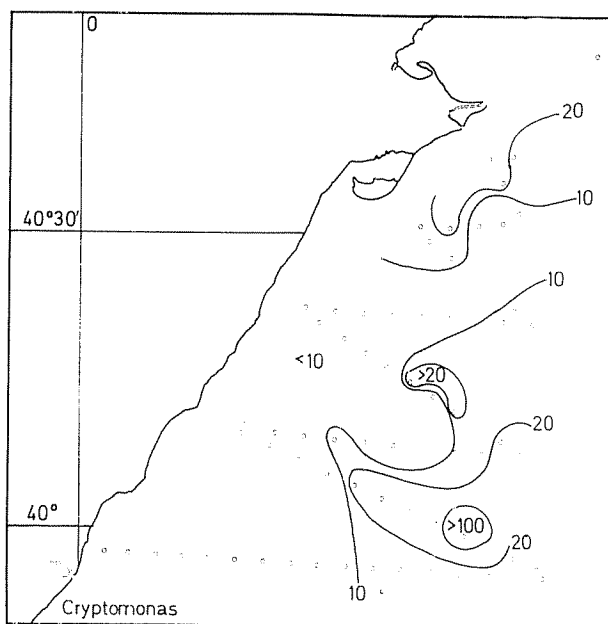


FIG. 4. — Distribution des deux espèces de *Cryptomonas* (cellules/ml).

le caractère d'affleurement de la zone, le nombre d'individus compté pour chaque échantillon était beaucoup plus élevé.

Pourtant l'examen des résultats de cette étude du phytoplancton de Castelló a donné des renseignements utiles. Pour tenter d'interpréter les composantes, on a commencé par faire des graphiques de la position

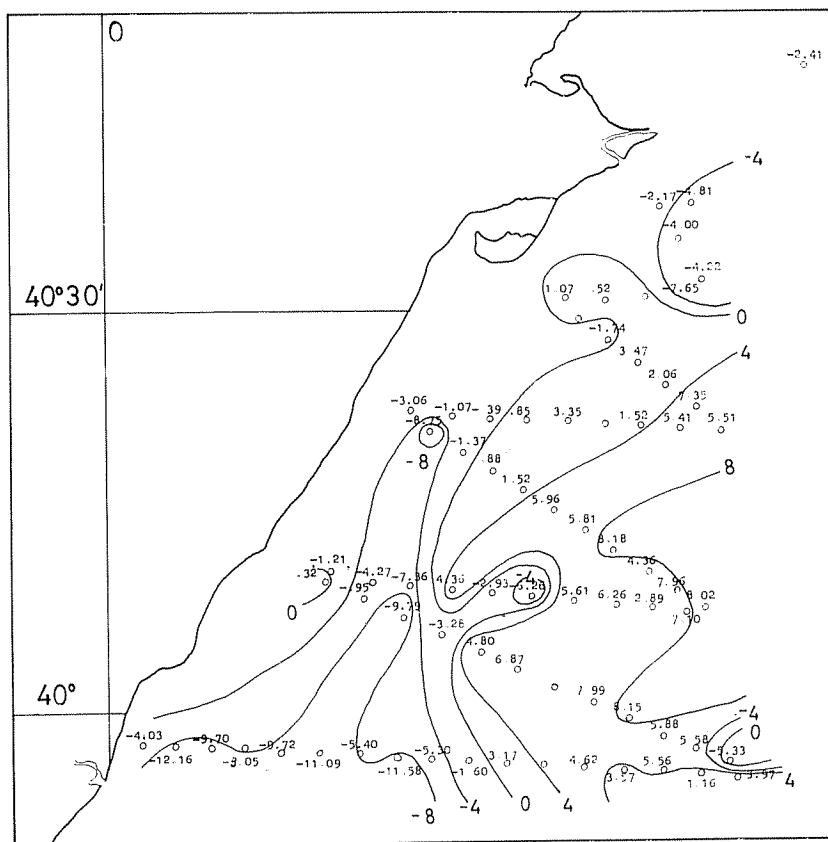


FIG. 5. — Distribution géographique des valeurs de la première composante.

des espèces dans l'espace des axes I-II, I-III, II-III, en utilisant comme coordonnées des points les coefficients de charge («factor loadings») correspondants (figs. 1, 2, 3).

Comme on peut voir, les espèces *Cryptomonas pseudobaltica* et *Cryptomonas sp.* se détachent clairement par la valeur élevée des coefficients de charge pour le premier axe. Par contre, la plupart des diatomées se signale par des valeurs négatives, quelquefois très fortes. Margalef avait

déjà cartographié la distribution des deux espèces de *Cryptomonas* (figure 4) et avait signalé la particularité de leur distribution.

On a calculé aussi la valeur des composantes pour chaque station, et on a cartographié les résultats pour les trois premières, en traçant des lignes reliant les points à la même valeur. Les figures 5, 6 et 7

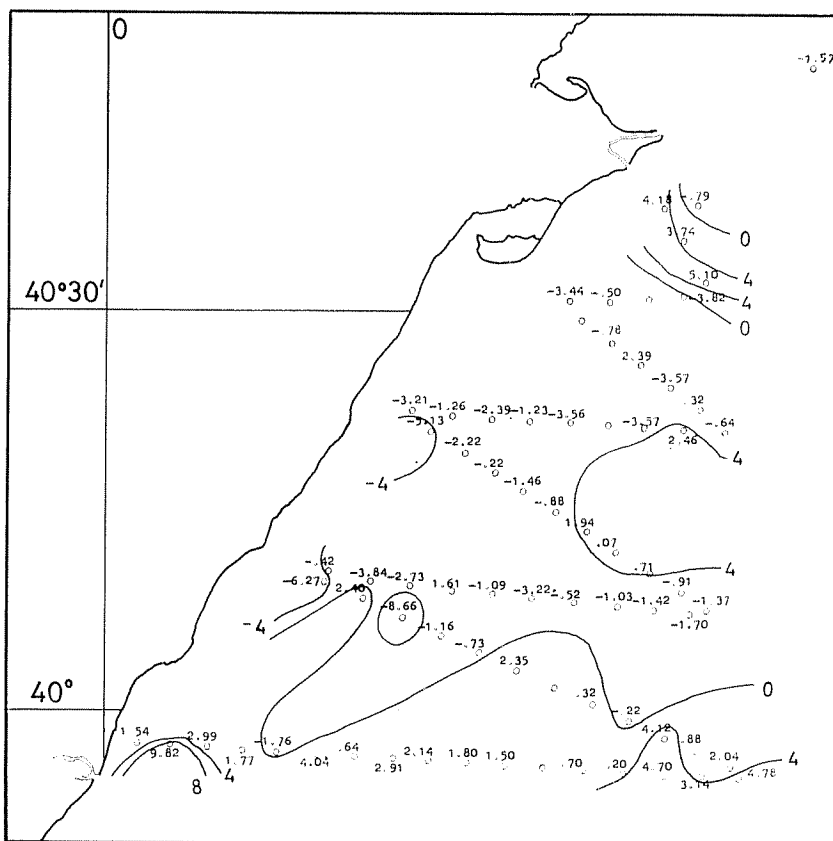


Fig. 6. — Distribution géographique des valeurs de la deuxième composante.

montrent la distribution des valeurs trouvées pour ces composantes. Il faut remarquer, qu'ici se pose un problème important : il doit y avoir une certaine contiguïté ou continuité entre les valeurs trouvées pour des stations voisines. On ne peut pas supposer une indépendance pour des résultats voisins dans l'espace (ou dans le temps). C'est un peu le problème inversé de la destruction des « patterns » dans l'espace ou dans le temps quand on cherche des relations entre des valeurs ponctuelles de variables. Une autre question intéressante, suggérée par Margalef, est

celle de savoir s'il existe quelque rapport généralisable entre les «pat-terns» géographiques relatifs à différentes composantes principales, étant donnée la condition d'orthogonalité de ces composantes.

La séparation relativement grande entre les stations, dans ce cas, et le caractère obligatoirement subjectif du dessin par l'observateur des courbes de niveau sont des inconvénients assez importants pour l'étude

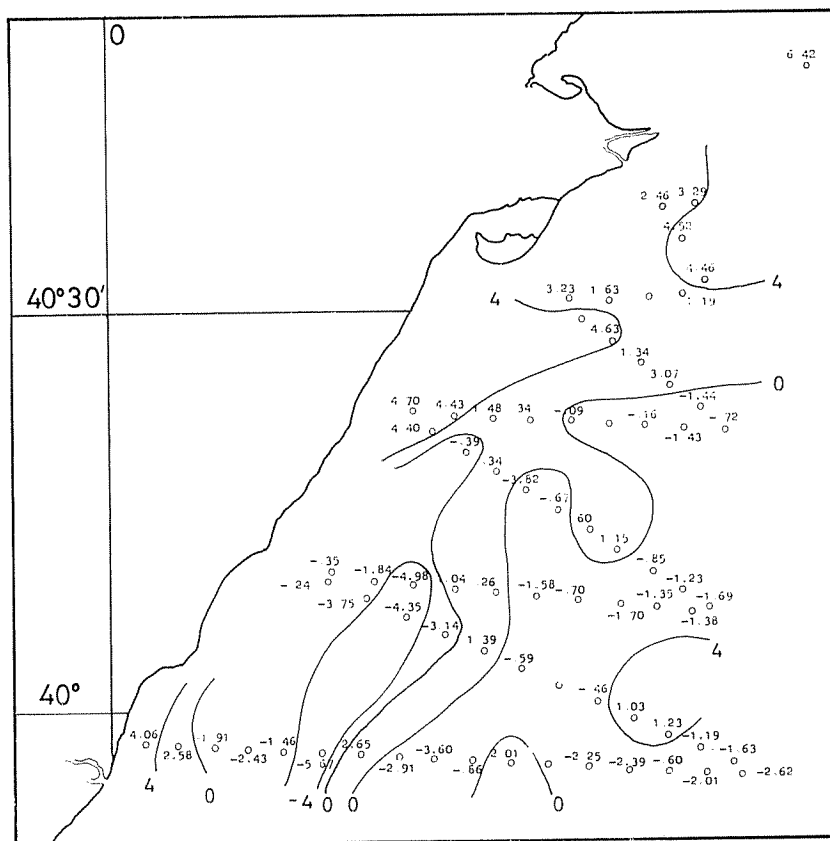


FIG. 7. — Distribution géographique des valeurs de la troisième composante.

de ces cartes. De toute façon, on peut voir que la distribution spatiale de la première composante rappelle un peu celle des *Cryptomonas*, comme d'ailleurs on pouvait s'attendre. Les distributions spatiales des 2^e et 3^e composantes sont difficiles à interpréter.

On a calculé aussi les corrélations des 4 premières composantes avec les données de fluorescence, chlorophylle, diversité spécifique (selon la formule $D = -\sum p_i \log_2 p_i$), indice de pigments DO_{430}/DO_{665} (DO : den-

sité optique), et les mesures chimiques dont on disposait. Les corrélations des composantes avec les concentrations de chlorophylle ont été calculées en utilisant uniquement les valeurs correspondants aux 33 stations où l'on avait déterminé le pigments.

On a représenté dans la table II les valeurs trouvés ; no notera ici seulement quelques aspects. La première composante présente des corrélations significatives de signe positif avec la fluorescence, et négatif avec phosphates, nitrates et diversité. La deuxième composante présente une corrélation positive avec le nombre total de cellules, et négative avec l'indice DO_{430}/DO_{665} .

Les calculs de corrélations entre des valeurs ponctuelles ont le défaut général de ne pas tenir compte de la structure spatiale ou temporelle des communautés étudiées. Comme Margalef a signalé dans ce même séminaire, on devrait chercher la façon de comparer des structures en bloc. Le calcul de surfaces de tendance et l'étude des distributions des résidus peut être un pas en ce sens.

TABLE II

Corrélation des composantes avec quelques paramètres chimiques et biologiques

Composantes	Fluor.	Diversité	log. Cel. tot.	Silicate	Phosphate	Nitrate	Chlorophyl.	$\frac{DO_{430}}{DO_{665}}$
1	0,28*	-0,67*	0,03	-0,14	-0,45*	-0,30*	0,28	0,07
2	0,08	0,04	0,35*	-0,02	-0,04	0,13	0,19	-0,42*
3	0,42*	0,16	0,37*	0,42*	0,53*	0,53*	0,37*	0,46*
4	0,16	0,02	0,13	0,26	0,16	0,40*	0,12	0,09

(* : Significatif, $p < 0,05$)

En conclusion, on peut dire que dans cet exemple de Castelló, l'analyse en composantes principales permet d'identifier une distribution opposée de la plupart des diatomées et des deux espèces de *Cryptomonas* trouvées. D'une part, la distribution des *Cryptomonas* est caractérisée par un centre de densité maximum dans l'angle SE de la zone étudiés ; de l'autre part, la distribution des diatomées (on considère le nombre total d'individus de ce groupe) présente un centre de densité maximum près du débouché de l'Ebre. Cette zone, avec une haute concentration de nutriments, possède une diversité relativement haute, qu'on peut attribuer aux conditions de mélange. Ça explique que la première composante soit, à la fois, en corrélation positive avec la diversité spécifique et avec phosphates et nitrates. On peut signaler encore une troisième zone, située à l'angle SW de l'espace parcouru, où l'on trouve aussi des valeurs relativement hauts de diversité, mais avec des concentrations de nutriments basses.

BIBLIOGRAPHIE

- BLASCO, D. — 1971. Composición y distribución del fitoplancton en la región del afloramiento de las costas peruanas. *Inv. Pesq.*, 35 : 61-112.
- MARGALEF, R. y GONZÁLEZ BERNÁLDEZ, F. — 1969. Grupos de especies asociadas en el fitoplancton del mar Caribe (NE de Venezuela). *Inv. Pesq.*, 33 : 287-312.

DISCUSSION

- IBANEZ. — Il serait peut être intéressant étant donné la proximité de l'Ebre de corréler les composantes avec Température et salinité et surtout la stabilité verticale σ_D (écart type des densités). On pourrait aussi étudier les masses d'eau du point de vue physico-chimique par l'analyse factorielle à l'aide des 67 stations et comparer les cartographies des groupes de paramètres (composantes climatiques) aux composantes que vous avez trouvées.